(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
28 November 2002 (28.11.2002)

PCT

(10) International Publication Number
WO 02/096128 A2

(51) International Patent Classification[7]: H04Q 3/00, 7/24

(21) International Application Number: PCT/IB02/02212

(22) International Filing Date: 2 April 2002 (02.04.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
0108041.5 30 March 2001 (30.03.2001) GB

(71) Applicant *(for all designated States except US)*: **NOKIA CORPORATION** [FI/FI]; Keilalahdentie 4, FIN-02150 ESPOO (FI).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **KISS, Krisztian** [HU/FI]; Saastajankatu 13C 14, FIN-33840 Tampere (FI). **ISOMAKI, Markus** [FI/FI]; Ajurinkatu 3B 43, FIN-02699 Espoo (FI). **PESSI, Pekka** [FI/FI]; Neitojenranta 1 A 9, FIN-00810 Helsinki (FI).

(74) Agents: **WILLIAMS, David, John** et al.; Page White & Farrer, 54 Doughty Street, London WC1N 2LS (GB).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with declaration under Article 17(2)(a); without abstract; title not checked by the International Searching Authority*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 02/096128 A2

(54) Title: PRESENCE SERVER IN IP MULTIMEDIA

(57) Abstract:

PRESENCE SERVER IN IP MULTIMEDIA

Field of the Invention

The present invention relates to the provision of a system architecture in an packet switched environment, and particularly to the implementation in such an architecture of means for providing information about a user's presence.

Background to the Invention

In third generation (3G) mobile networks services are provided over IP networks, which results in the integration of voice and data applications. One of the major candidates for the emerging new IP based services is to provide information about the user's presence. Presence is defined as subscription to and notification of changes in the communications state of a user. This communications state consists of the set of: communications means; communications address; and status of that user.

In third generation networks, call control and the service creation environment are based on a session initiation protocol (SIP), as described in 3rd Generation Partnership Project, Technical Specification Group Services and System Aspects, IP Multi-Media Sub-System - Stage 2, 3G TS 23.228 version 1.7.0, February 2001.

Internet Engineering Task Force, Internet Draft, draft-rosenberg-impp-presence-01.txt, Rosenberg et al, published 2nd March 2001, the contents of which are incorporated herein by reference as Annex A, proposes an extension to SIP for subscriptions and notifications of user presence. User presence is defined as the willingness and ability of a user to communicate with other users on the network. Historically, presence has been limited to "on-line" and "off-line" indicators. The notion of presence in Rosenberg et al is broader. Subscriptions and notifications of user presence are

supported by defining an event package within the general SIP event notification framework. This protocol is also compliant with the common presence and instant messaging (CPIM) framework. However, such proposal does not include any consideration of multimedia environments.

The SIP extension defined in Rosenberg et al is based on the concept of a presence agent (PA), which is a new logical entity that is capable of accepting subscriptions (through a SUBSCRIBE message), storing a subscription state, and generating notifications (through a NOTIFY message) when there are changes in user presence.

The aim of this invention is to provide a technique for the session initiation protocol registration, subscription and notification procedures in an internet protocol multimedia subsystem.

Summary of the Invention

The aim of the present invention is achieved by providing a presence server in the architecture. The presence server is preferably provided as part of the application and services 'cloud' or environment. By providing the presence server in the architecture, subscribers are able to receive information about other subscriber's presence.

The present invention is related to 3GPP (3rd Generation Partnership Project) Release 5/6 standardization.

In accordance with the present invention there is provided a packet switched environment, including the functionality of a presence server in an application and services environment.

The packet switched environment is preferably an internet protocol multimedia environment, and preferably a subsystem of an all-IP telecommunications network.

In a first embodiment the interrogating call state control function (I-CSCF) updates the presence information in the presence server by forking an incoming REGISTER message.

In a second embodiment the serving call state control function (S-CSCF) acts as a presence agent and the presence server provides the task of storing information about the subscribers.

In a third embodiment the presence server in the internet protocol multimedia (IM) sub-system behaves as a presence agent and the serving call state control function (S-CSCF) uses a separate REGISTER transaction to update the presence information in the presence server.

The invention thus solves the problem of routing of the REGISTER, SUBSCRIBE and NOTIFY messages in an internet protocol multimedia (IM) subsystem.

Brief Description of the Drawings

The invention will now be described by way of reference to the accompanying drawings, in which:

Figure 1 represents the 3GPP Release 5 architecture embodying the present invention;

Figure 2 represents the flow of REGISTER messages in a first embodiment of the present invention;

Figure 3 represents the flow of SUBSCRIBE messages in a first embodiment of the present invention;

Figure 4 represents the flow of NOTIFY messages in a first embodiment of the present invention;

Figure 5 represents the flow of SUBSCRIBE messages in a second embodiment of the present invention;

Figure 6 represents the flow of NOTIFY messages in a second embodiment of the present invention;

Figure 7 represents the flow of REGISTER messages in a third embodiment of the present invention;

Figure 8 represents the flow of SUBSCRIBE messages in a third embodiment of the present invention; and

Figure 9 represents the flow of NOTIFY messages in a third embodiment of the present invention.

Description of Preferred Embodiments

The present invention places a presence server in the internet multimedia subsystem of the 3GPP Release 5 architecture as part of the 'application and services cloud'. The internet protocol multimedia subsystem refers to the set of Core Network entities using the services provided by the packet switched domain of the 3GPP Release 5 architecture to offer multimedia services. The entities of the internet protocol multimedia subsystem are the CSCF, the MGCF, the MRF and some adaptation entities. The representation of the extended architecture is shown in Figure 1.

Figure 1 is based on a basic 3GPP architecture in accordance with the architecture defined in 3rd Generation Partnership Project; Technical Specification Group Services and Systems Aspects; Architecture for an All-IP Network; 3G TR 23.922 version 1.0.0; October 1999, the contents of which are herein incorporated by reference as Annex B. However, the architecture disclosed therein is modified, as shown in Figure 1, to include a presence server in accordance with the present invention.

The present invention is particularly concerned with the flow of REGISTER, SUBSCRIBE and NOTIFY messages in a 3GPP network.

The present invention will now be described in further detail with reference to three exemplary embodiments of REGISTER, SUBSCRIBE, and NOTIFY message flows. It should be noted that

only the necessary parts of the network - and message flows - needed for operation of the present invention are described in the following examples.

One general requirement for all the described embodiments is that the user's profile information must contain the name of the presence server associated with that user.

The routing of the REGISTER/SUBSCRIBE/NOTIFY messages of a first embodiment is described hereinbelow with reference to Figures 2 to 4.

In this first embodiment, the interrogating call state control function (I-CSCF) updates the presence information in the presence server by forking an incoming REGISTER message.

A first network corresponds to the visited network of the presence user agent (PUA), and includes user equipment (UE) 200, and a proxy call state control function (P-CSCF) 202. The first network is also the presence agent's network.

A second network corresponds to the home network of the presence user agent (PUA), and includes an interrogating call state control function (I-CSCF) 204, a (HSS) 206, a serving call state control function (S-CSCF) 208, and a presence server (PS) 210.

A third network corresponds to the home network of the subscriber, and includes a UE subscriber 212, a first proxy call state control function (P-CSCF#1) 214, a first serving call state control function (S-CSCF#1) 216, an interrogating call state control function (I-CSCF) 218, a (HSS) 220, a second serving call-state control function (S-CSCF#2) 222, and a second proxy call state control function (P-CSCF#2) 224.

Referring to Figure 2, there is illustrated an extended registration message flow in accordance with this first embodiment of the present invention.

As represented by step 230, the routing of the REGISTER message, initiated by the user equipment 200, takes place between the UE 200, the first network P-CSCF 202 and the second network I-CSCF 204. The name of the presence server 210 is part of the subscriber's profile, and this is retrieved by the I-CSCF 204 from the HSS 206 in a step 232. In a step 234 the I-CSCF 204 selects a S-CSCF for the session initiation, which in this example is the S-CSCF 208.

As represented by messages 236 and 238, the I-CSCF 204 forks the incoming REGISTER message such that, in accordance with this embodiment, it is forwarded to both the S-CSCF 208 and the PS 210. Thereafter "200 OK" messages are transmitted back to the UE 200 along the reverse route.

Referring to Figure 3, there is illustrated a routing of the SUBSCRIBE message flow in accordance with this first embodiment of the present invention.

As represented by messages 240, 242, and 244, the SUBSCRIBE message is routed to the I-CSCF 204 from the UE subscriber 212 via the P-CSCF#1 214 and S-CSCF#1 216.

As represented by message 246, the I-CSCF 204 routes the received SUBSCRIBE message directly to the PS 210. Thereafter "202 Accepted" messages are routed back to the UE subscriber along the reverse route.

Referring to Figure 4, there is illustrated a routing of the NOTIFY message flow from the presence server 210 in accordance with this first embodiment of the present invention.

The PS 210, as represented by the message 248, forwards the NOTIFY message to the I-CSCF 218. Following a local query to the HSS 220 in step 249 the I-CSCF, as represented by messages 250, 252 and 254, forwards the NOTIFY message to the UE subscriber 226 via the S-CSCF#2 222 and the P-CSCF#2 224. Thus the PS 210 sends the NOTIFY message directly to the other

networks I-CSCF 218. Once again, "200 OK" messages are routed back to the PS 210 along the reverse route.

In summary, the first embodiment sets the following new requirements to the network elements placed in the architecture:

1. I-CSCF downloads (part of) the subscriber's profile from HSS in order to find the subscriber's presence server;
2. I-CSCF forks the incoming REGISTER of PUA to the S-CSCF and to the presence server;
3. I-CSCF routes the incoming SUBSCRIBE requests to the presence server directly; and
4. The presence server sends the outgoing NOTIFYs directly to other network's I-CSCF.

The routing of the SUBSCRIBE/NOTIFY messages in accordance with a second embodiment of the invention is described hereinbelow with reference to Figures 5 and 6.

Since the S-CSCF stores the subscriber profile and the contact information provided in the REGISTER message, a trivial solution is for the S-CSCF to act as a presence agent. One possible problem with this solution is that the S-CSCF would have to store information about all the subscribers and generate NOTIFY messages to all of them.

Therefore, in the second described embodiment, the task of storing information about each subscribers is provided by the newly introduced presence server. For the S-CSCF it is enough to receive only the first SUBSCRIBE message for the presentity, since it will generate the NOTIFY message for the one subscription, and this NOTIFY message will be forked to the subscribers by the proxy server, which has the information about the subscribers.

The registration flow in this solution corresponds to the normal registration as defined by 3G TS 23.228 version 1.7.0, February 2001, discussed hereinabove in the introduction.

The routing of the SUBSCRIBE/NOTIFY messages in accordance with the second embodiment of the present invention is described hereinbelow with reference to Figures 5 and 6. Elements of the first, second and third networks corresponding to elements shown in Figures 2 to 4 are referenced in Figures 5 and 6 using the same reference numerals.

In addition to the elements shown in Figures 2 to 4, for the purposes of describing the second embodiment the third network, corresponding to the home network of the subscriber, includes two user equipment subscribers UE subs#1 302 and UE subs#2 304. Furthermore, the second network, corresponding to the home network of the PUA, includes a second serving call state control function S-CSCF#2 306.

Referring to Figure 5, there is illustrated a routing of the SUBSCRIBE message flow in accordance with this embodiment of the present invention.

A SUBSCRIBE message from the first user equipment subscriber 302 is forwarded to the P-CSCF#1 214, as illustrated by message 308a. Such message is forwarded, in turn, to the S-CSCF#1 216 and the I-CSCF 204 as illustrated by messages 310a and 312a. Following a local query in step 313a, the SUBSCRIBE message is forwarded to the PS 210, as represented by message 314a. Subsequent thereto, the SUBSCRIBE message is forwarded from the PS 210 to the S-CSCF#2 306, as represented by message 316.

Responsive to the SUBSCRIBE message 316 from the presence server 210 corresponding to the first user equipment subscriber 302 of the third network, the S-CSCF#2 returns an

acknowledgement message by way of an accept signal, as represented by arrow 318.

Similarly a SUBSCRIBE message from the second user equipment subscriber 304 is forwarded to the P-CSCF#1 214, as illustrated by message 308b. Such message is forwarded, in turn, to the S-CSCF#1 216 and the I-CSCF 204 as illustrated by messages 310b and 312b. Following a local query in step 313b, the SUBSCRIBE message is forwarded to the PS 210, as represented by message 314b.

It is not necessary for the presence server 210 to forward any subsequent SUBSCRIBE messages from user equipment of the third network, as the S-CSCF#2 306 has already provided a successful acknowledgement.

Acceptance signals for each user subscriber are returned to the respective subscribers along reverse paths, responsive to receipt of the message 318 and an appropriate SUBSCRIBE message 314.

Referring to Figure 6, there is illustrated a routing of the NOTIFY message flow from the presence server 210 in accordance with this embodiment of the present invention.

As represented by message 320, a single NOTIFY message is forwarded from the S-CSCF#1 to the presence server 210.

The presence server, as represented by message 322a then forwards a first NOTIFY message to the I-CSCF 218. Following a first local query 324a, the first NOTFIY message is forwarded to the first user equipment subscriber 302 via, in turn, the S-CSCF#2 222 and the P-CSCF#2 224, as represented by messages 326a, 328a, and 330a.

The presence server, as represented by message 322b also forwards a second NOTIFY message to the I-CSCF 218. Following a second local query 324b, the second NOTFIY message is

forwarded to the first user equipment subscriber 302 via, in turn, the S-CSCF#2 222 and the P-CSCF#2 224, as represented by messages 326b, 328b, and 330b.

"200 OK" messages are returned to the presence server 210 via a reverse path, and a single "OK" message forwarded to the S-CSCF#1.

The routing of the REGISTER/SUBSCRIBE/NOTIFY messages in accordance with a third embodiment of the present invention is described hereinbelow with reference to Figures 7 to 9.

The solution described in this third embodiment can be summarised as: the presence server in the internet protocol multimedia subsystem behaves as a presence agent, and the S-CSCF uses a separate REGISTER transaction to update the presence information in the presence server.

The routing of the REGISTER/SUBSCRIBE/NOTIFY methods is described hereinafter with reference to Figures 7 to 9. Elements of the first, second and third networks corresponding to elements shown in Figures 2 to 6 are referenced in Figures 7 to 9 using the same reference numerals.

Referring to Figure 7, there is illustrated an extended registration message flow in accordance with this third embodiment of the present invention.

As represented by step 230, the routing of the REGISTER message takes place between the UE 200, the first network P-CSCF 202 and the second network I-CSCF 204.

Thereafter, in a step 702, the I-CSCF 204 selects a S-CSCF 208 for the session initiation, which in this example is the S-CSCF 208.

As represented by message 704, the REGISTER message is then forwarded to the S-CSCF 208.

The name of the presence server 210 is part of the subscriber's profile, and this is retrieved by the I-CSCF 204 from the HSS 206 in a step 708.

Following successful initiation with the S-CSCF 208, a "200 OK" message is transmitted to the UE 200 along a reverse path. Thereafter, as represented by message 710, the REGISTER message is forwarded to the presence server 210. The message 710 constitutes a separate REGSITER transaction with which the S-CSCF updates the presence information in the PS.

If the presence update fails at the presence server, the S-CSCF generates a notification to the UE 200 indicating the presence update failure event. For this notification a SIP NOTIFY message can be used, for example, containing an event header with a new presence failure reason code. This example is illustrated in Figure 7 labelled as 711. The NOTIFY is then acknowledged by the UE 200 to the S-CSCF 208 using a "200 OK" message.

Referring to Figure 8, there is illustrated a routing of the SUBSCRIBE message flow in accordance with this third embodiment of the present invention.

A SUBSCRIBE message from the user equipment subscriber 212 is forwarded to the P-CSCF#1 214, as illustrated by message 712. Such message is forwarded, in turn, to the S-CSCF#1 216 and the I-CSCF 204 as illustrated by messages 714 and 716. Following a local query in step 718, the SUBSCRIBE message is forwarded to the S-CSCF#2 306, as represented by message 720. Subsequent thereto, the SUBSCRIBE message is forwarded from the S-CSCF#2 306 to the PS 210, as represented by message 720.

Thereafter a "202 Accepted" message is sent along the reverse path.

Referring to Figure 9, there is illustrated a routing of the NOTIFY message flow from the presence server 210 in accordance with this third embodiment of the present invention.

As represented by message 722, a NOTIFY message is forwarded from the presence server 210 to the S-CSCF#1 306. The S-CSCF#1 306, as represented by message 724 then forwards the NOTIFY message to the I-CSCF 218. Following a local query in step 726, the NOTIFY message is forwarded to the user equipment subscriber 226 via, in turn, the S-CSCF#2 222 and the P-CSCF#2 224, as represented by messages 728, 730, and 732.

Thereafter a "200 OK" message is sent along a reverse path.

The new requirements for the network elements placed in the architecture due to the third embodiment cab be summarised as below:

1. The S-CSCF updates the presence information in the PS with a separate REGISTER transaction.
2. If the presence update fails at the PS, the S-CSCF generates a notification message (e.g. SIP NOTIFY using the Evenyt header with a new presence failure reason code) to the UE indicating the presence failure update event.
3. The SUBSCRIBE generated by the subscriber has to be routed by the network as it would be a normal INVITE, only the S-CSCF of the PUA routes the SUBSCRIBE to the PS associated with the presentity.
4. The NOTIFY generated by the PS has to be routed by the network as it would be a normal INVITE.

Although the present invention has been described with reference to three exemplary embodiments, the present invention more generally presents ways of implementing the

idea presented in Rosenberg et al with 3GPP proposed architecture.

Internet Engineering Task Force
Internet Draft
draft-rosenberg-impp-presence-01.txt
March 2, 2001
Expires: September 2001

SIMPLE WG
Rosenberg et al.
Various Places

SIP Extensions for Presence

STATUS OF THIS MEMO

Abstract

This document proposes an extension to SIP for subscriptions and notifications of user presence. User presence is defined as the willingness and ability of a user to communicate with other users on the network. Historically, presence has been limited to "on-line" and "off-line" indicators; the notion of presence here is broader. Subscriptions and notifications of user presence are supported by defining an event package within the general SIP event notification framework. This protocol is also compliant with the Common Presence and Instant Messaging (CPIM) framework.

1 Introduction

Presence is (indirectly) defined in RFC2778 [1] as subscription to and notification of changes in the communications state of a user.

This communications state consists of the set of communications means, communications address, and status of that user. A presence protocol is a protocol for providing such a service over the Internet or any IP network.

This document proposes an extension to the Session Initiation Protocol (SIP) [2] for presence. This extension is a concrete instantiation of the general event notification framework defined for SIP [3], and as such, makes use of the SUBSCRIBE and NOTIFY methods defined there. User presence is particularly well suited for SIP. SIP registrars and location services already hold user presence information; it is uploaded to these devices through REGISTER messages, and used to route calls to those users. Furthermore, SIP networks already route INVITE messages from any user on the network to the proxy that holds the registration state for a user. As this state is user presence, those SIP networks can also allow SUBSCRIBE requests to be routed to the same proxy. This means that SIP networks can be reused to establish global connectivity for presence subscriptions and notifications.

This extension is based on the concept of a presence agent, which is a new logical entity that is capable of accepting subscriptions, storing subscription state, and generating notifications when there are changes in user presence. The entity is defined as a logical one, since it is generally co-resident with another entity, and can even move around during the lifetime of a subscription.

This extension is also compliant with the Common Presence and Instant Messaging (CPIM) framework that has been defined in [4]. This allows SIP for presence to easily interwork with other presence systems compliant to CPIM.

2 Definitions

This document uses the terms as defined in [1]. Additionally, the following terms are defined and/or additionally clarified:

Presence User Agent (PUA): A Presence User Agent manipulates presence information for a presentity. In SIP terms, this means that a PUA generates REGISTER requests, conveying some kind of information about the presentity. We explicitly allow multiple PUAs per presentity. This means that a user can have many devices (such as a cell phone and PDA), each of which is independently generating a component of the overall presence information for a presentity. PUAs push data into the presence system, but are outside of it, in that they do not receive SUBSCRIBE messages, or send NOTIFY.

Presence Agent (PA): A presence agent is a SIP user agent which is capable of receiving SUBSCRIBE requests, responding to them, and generating notifications of changes in presence state. A presence agent must have complete knowledge of the presence state of a presentity. Typically, this is accomplished by co-locating the PA with the proxy/registrar, or the presence user agent of the presentity. A PA is always addressable with a SIP URL.

Presence Server: A presence server is a logical entity that can act as either a presence agent or as a proxy server for SUBSCRIBE requests. When acting as a PA, it is aware of the presence information of the presentity through some protocol means. This protocol means can be SIP REGISTER requests, but other mechanisms are allowed. When acting as a proxy, the SUBSCRIBE requests are proxied to another entity that may act as a PA.

Presence Client: A presence client is a presence agent that is colocated with a PUA. It is aware of the presence information of the presentity because it is co-located with the entity that manipulates this presence information.

3 Overview of Operation

In this section, we present an overview of the operation of this extension.

When an entity, the subscriber, wishes to learn about presence information from some user, it creates a SUBSCRIBE request. This request identifies the desired presentity in the request URI, using either a presence URL or a SIP URL. The subscription is carried along SIP proxies as any other INVITE would be. It eventually arrives at a presence server, which can either terminate the subscription (in which case it acts as the presence agent for the presentity), or proxy it on to a presence client. If the presence client handles the subscription, it is effectively acting as the presence agent for the presentity. The decision about whether to proxy or terminate the SUBSCRIBE is a local matter; however, we describe one way to effect such a configuration, using REGISTER.

The presence agent (whether in the presence server or presence client) first authenticates the subscription, then authorizes it. The means for authorization are outside the scope of this protocol, and we expect that many mechanisms will be used. Once authorized, the presence agent sends a 202 Accepted response. It also sends an immediate NOTIFY message containing the state of the presentity. As the state of the presentity changes, the PA generates NOTIFYs for all

subscribers.

The SUBSCRIBE message effectively establishes a session with the presence agent. As a result, the SUBSCRIBE can be record-routed, and rules for tag handling and Contact processing mirror those for INVITE. Similarly, the NOTIFY message is handled in much the same way a re-INVITE within a call leg is handled.

4 Naming

A presentity is identified in the most general way through a presence URI [4], which is of the form pres:user@domain. These URIs are protocol independent. Through a variety of means, these URIs can be resolved to determine a specific protocol that can be used to access the presentity. Once such a resolution has taken place, the presentity can be addressed with a sip URL of nearly identical form: sip:user@domain. The protocol independent form (the pres: URL) can be thought of as an abstract name, akin to a URN, which is used to identify elements in a presence system. These are resolved to concrete URLs that can be used to directly locate those entities on the network.

When subscribing to a presentity, the subscription can be addressed using the protocol independent form or the sip URL form. In the SIP context, "addressed" refers to the request URI. It is RECOMMENDED that if the entity sending a SUBSCRIBE is capable of resolving the protocol independent form to the SIP form, this resolution is done before sending the request. However, if the entity is incapable of doing this translation, the protocol independent form is used in the request URI. Performing the translation as early as possible means that these requests can be routed by SIP proxies that are not aware of the presence namespace..

The result of this naming scheme is that a SUBSCRIBE request is addressed to a user the exact same way an INVITE request would be addressed. This means that the SIP network will route these messages along the same path an INVITE would travel. One of these entities along the path may act as a PA for the subscription. Typically, this will either be the presence server (which is the proxy/registrar where that user is registered), or the presence client (which is one of the user agents associated with that presentity).

SUBSCRIBE messages also contain logical identifiers that define the originator and recipient of the subscription (the To and From header fields). Since these identifiers are logical ones, it is RECOMMENDED that these use the protocol independent format whenever possible. This also makes it easier to interwork with other systems which recognize these forms.

The Contact, Record-Route and Route fields do not identify logical entities, but rather concrete ones used for SIP messaging. As such, they MUST use the SIP URL forms in both SUBSCRIBE and NOTIFY.

5 Presence Event Package

The SIP event framework [3] defines an abstract SIP extension for subscribing to, and receiving notifications of, events. It leaves the definition of many additional aspects of these events to concrete extensions, also known as event packages. This extension qualifies as an event package. This section fills in the information required by [3].

5.1 Package Name

The name of this package is "presence". This name MUST appear within the Event header in SUBSCRIBE request and NOTIFY request. This section also serves as the IANA registration for the event package "presence".

TODO: Define IANA template in sub-notify and fill it in here.

Example:

Event: presence

5.2 SUBSCRIBE bodies

The body of a SUBSCRIBE request MAY contain a body. The purpose of the body depends on its type. In general, subscriptions will normally not contain bodies. The request URI, which identifies the presentity, combined with the event package name, are sufficient for user presence.

We anticipate that document formats could be defined to act as filters for subscriptions. These filters would indicate certain user presence events that would generate notifies, or restrict the set of data returned in NOTIFY requests. For example, a presence filter might specify that the notifications should only be generated when the status of the users instant message inbox changes. It might also say that the content of these notifications should only contain the IM related information.

5.3 Expiration

User presence changes as a result of events that include:

- o Turning on and off of a cell phone
- o Modifying the registration from a softphone
- o Changing the status on an instant messaging tool

These events are usually triggered by human intervention, and occur with a frequency on the order of minutes or hours. As such, it is subscriptions should have an expiration in the middle of this range, which is roughly one hour. Therefore, the default expiration time for subscriptions within this package is 3600 seconds. As per [3], the subscriber MAY include an alternate expiration time. Whatever the indicated expiration time, the server MAY reduce it but MUST NOT increase it.

5.4 NOTIFY Bodies

The body of the notification contains a presence document. This document describes the user presence of the presentity that was subscribed to. All subscribers MUST support the presence data format described in [fill in with IMPP document TBD], and MUST list its MIME type, [fill in with MIME type] in an Accept header present in the SUBSCRIBE request.

Other presence data formats might be defined in the future. In that case, the subscriptions MAY indicate support for other presence formats. However, they MUST always support and list [fill in with MIME type of IMPP presence document] as an allowed format.

Of course, the notifications generated by the presence agent MUST be in one of the formats specified in the Accept header in the SUBSCRIBE request.

5.5 Processing Requirements at the PA

User presence is highly sensitive information. Because the implications of divulging presence information can be severe, strong requirements are imposed on the PA regarding subscription processing, especially related to authentication and authorization.

A presence agent MUST authenticate all subscription requests. This authentication can be done using any of the mechanisms defined for SIP. It is not considered sufficient for the authentication to be transitive; that is, the authentication SHOULD use an end-to-end mechanism. The SIP basic authentication mechanism MUST NOT be used.

It is RECOMMENDED that any subscriptions that are not authenticated do not cause state to be established in the PA. This can be accomplished by generating a 401 in response to the SUBSCRIBE, and then discarding all state for that transaction. Retransmissions of the SUBSCRIBE generate the same response, guaranteeing reliability even over UDP.

Furthermore, a PA MUST NOT accept a subscription unless authorization has been provided by the presentity. The means by which authorization are provided are outside the scope of this document. Authorization may have been provided ahead of time through access lists, perhaps specified in a web page. Authorization may have been provided by means of uploading of some kind of standardized access control list document. Back end authorization servers, such as a DIAMETER [5], RADIUS [6], or COPS [7], can also be used. It is also useful to be able to query the user for authorization following the receipt of a subscription request for which no authorization information was present. Appendix A provides a possible solution for such a scenario.

The result of the authorization decision by the server will be reject, accept, or pending. Pending occurs when the server cannot obtain authorization at this time, and may be able to do so at a later time, when the presentity becomes available.

Unfortunately, if the server informs the subscriber that the subscription is pending, this will divulge information about the presentity - namely, that they have not granted authorization and are not available to give it at this time. Therefore, a PA SHOULD generate the same response for both pending and accepted subscriptions. This response SHOULD be a 202 Accepted response.

If the server informs the subscriber that the subscription is rejected, this also divulges information about the presentity - namely, that they have explicitly blocked the subscription previously, or are available at this time and chose to decline the subscription. If the policy of the server is not to divulge this information, the PA MAY respond with a 202 Accepted response even though the subscription is rejected. Alternatively, if the policy of the presentity or the PA is that it is acceptable to inform the subscriber of the rejection, a 603 Decline SHOULD be used.

Note that since the response to a subscription does not contain any useful information about the presentity, privacy and integrity of SUBSCRIBE responses is not deemed important.

5.6 Generation of Notifications

Upon acceptance of a subscription, the PA SHOULD generate an

immediate NOTIFY with the current presence state of the presentity.

If a subscription is received, and is marked as pending or was rejected, the PA SHOULD generate an immediate NOTIFY. This NOTIFY should contain a valid state for the presentity, yet be one which provides no useful information about the presentity. An example of this is to provide an IM URL that is the same form as the presence URL, and mark that IM address as "not available". The reason for this process of "lying" is that without it, a subscriber could tell the difference between a pending subscription and an accepted subscription based on the existence and content of an immediate NOTIFY. The approach defined here ensures that the presence delivered in a NOTIFY generated by a pending or rejected subscription is also a valid one that could have been delivered in a NOTIFY generated by an accepted subscription.

If the policy of the presence server or the presentity is that it is acceptable to divulge information about whether the subscription succeeded or not, the immediate NOTIFY need not be sent for pending or rejected subscriptions.

Of course, once a subscription is accepted, the PA SHOULD generate a NOTIFY for the subscription when it determines that the presence state of the presentity has changed. Section 6 describes how the PA makes this determination.

For reasons of privacy, it will frequently be necessary to encrypt the contents of the notifications. This can be accomplished using the standard SIP encryption mechanisms. The encryption should be performed using the key of the subscriber as identified in the From field of the SUBSCRIBE. Similarly, integrity of the notifications is important to subscribers. As such, the contents of the notifications SHOULD be authenticated using one of the standardized SIP mechanisms. Since the NOTIFY are generated by the presence server, which may not have access to the key of the user represented by the presentity, it will frequently be the case that the NOTIFY are signed by a third party. It is RECOMMENDED that the signature be by an authority over domain of the presentity. In other words, for a user pres:user@example.com, the signator of the NOTIFY SHOULD be the authority for example.com.

5.7 Rate Limitations on NOTIFY

For reasons of congestion control, it is important that the rate of notifications not become excessive. As a result, it is RECOMMENDED that the PA not generate notifications for a single presentity at a rate faster than once every 5 seconds.

5.8 Refresh Behavior

Since SUBSCRIBE is routed by proxies as any other method, it is possible that a subscription might fork. The result is that it might arrive at multiple devices which are configured to act as a PA for the same presentity. Each of these will respond with a 202 response to the SUBSCRIBE. Based on the forking rules in SIP, only one of these responses is passed to the subscriber. However, the subscriber will receive notifications from each of those PA which accepted the subscriptions. The SIP event framework allows each package to define the handling for this case.

The processing in this case is identical to the way INVITE would be handled. The 202 Accepted to the SUBSCRIBE will result in the installation of subscription state in the subscriber. The subscription is associated with the To and From (both with tags) and Call-ID from the 202. When notifications arrive, those from the PA's whose 202's were discarded in the forking proxy will not match the subscription ID stored at the subscriber (the From tags will differ). These SHOULD be responded to with a 481. This will disable the subscriptions from those PA. Furthermore, when refreshing the subscription, the refresh SHOULD make use of the tags from the 202 and make use of any Contact or Record-Route headers in order to deliver the SUBSCRIBE back to the same PA that sent the 202.

The result of this is that a presentity can have multiple PAs active, but these should be homogeneous, so that each can generate the same set of notifications for the presentity. Supporting heterogeneous PAs, each of which generated notifications for a subset of the presence data, is complex and difficult to manage. If such a feature is needed, it can be accomplished with a B2BUA rather than through a forking proxy.

6 Publication

The user presence for a presentity can be obtained from any number of different ways. Baseline SIP defines a method that is used by all SIP clients - the REGISTER method. This method allows a UA to inform a SIP network of its current communications addresses (ie., Contact addresses) . Furthermore, multiple UA can independently register Contact addresses for the same SIP URL. These Contact addresses can be SIP URLs, or they can be any other valid URL.

Using the register information for presence is straightforward. The address of record in the REGISTER (the To field) identifies the presentity. The Contact headers define communications addresses that describe the state of the presentity. The use of the SIP caller preferences extension [8] is RECOMMENDED for use with UAs that are

interested in presence. It provides additional information about the Contact addresses that can be used to construct a richer presence document. The "description" attribute of the Contact header is explicitly defined here to be used as a free-form field that allows a user to define the status of the presentity at that communications address.

We also allow REGISTER requests to contain presence documents, so that the PUAs can publish more complex information.

Note that we do not provide for locking mechanisms, which would allow a client to lock presence state, fetch it, and update it atomically. We believe that this is not neeeded for the majority of use cases, and introduces substantial complexity. Most presence operations do not require get-before-set, since the SIP register mechanism works in such a way that data can be updated without a get.

The application of registered contacts to presence increases the requirements for authenticity. Therefore, REGISTER requests used by presence user agents SHOULD be authenticated using either SIP authentication mechanisms, or a hop by hop mechanism.

To indicate presence for instant messaging, the UA MAY either register contact addresses that are SIP URLs with the "methods" parameter set to indicate the method MESSAGE, or it MAY register an IM URL.

TODO: This section needs work. Need to define a concrete example of mapping a register to a presence document, once IMPP generates the document format.

6.1 Migrating the PA Function

It is important to realize that the PA function can be colocated with several elements:

- It can be co-located with the proxy server handling registrations for the presentity. In this way, the PA knows the presence of the user through registrations.
- It can be co-located with a PUA for that presentity. In the case of a single PUA per presentity, the PUA knows the state of the presentity by sheer nature of its co-location.
- It can be co-located in any proxy along the call setup path. That proxy can learn the presence state of the presentity by generating its own SUBSCRIBE in order to determine it. In this case, the PA is effectively a B2BUA.

Because of the soft-state nature of the subscriptions, it becomes possible for the PA function to migrate during the lifetime of a subscription. The most workable scenario is for the PA function to migrate from the presence server to the PUA, and back.

Consider a subscription that is installed in a presence server. Assume for the moment that the presence server can determine that a downstream UA is capable of acting as a PA for the presentity. When a subscription refresh arrives, the PA destroys its subscription, and then acts as a proxy for the subscription. The subscription is then routed to the UA, where it can be accepted. The result is that the subscription becomes installed in the PUA.

For this migration to work, the PUA MUST be prepared to accept SUBSCRIBE requests which already contain tags in the To field. Furthermore, the PUA MUST insert a Contact header into the 202, and this header MUST be used by the subscriber to update the contact address for the subscription.

TODO: Does this work? What about getting a Record-Route in place at the PUA. This might only be possible for refreshes that don't use Route or tags.

The presence server determines that a PUA is capable of supporting a PA function through the REGISTER message. Specifically, if a PUA wishes to indicate support for the PA function, it SHOULD include a contact address in its registration with a caller preferences "methods" parameter listing SUBSCRIBE.

7 Mapping to CPIM

This section defines how a SIP for presence messages are converted to CPIM, and how a CPIM messages are converted to SIP for presence. SIP to CPIM conversion occurs when a SIP system sends a SUBSCRIBE request that contains a pres URL or SIP URL that corresponds to a user in a domain that runs a different presence protocol. CPIM to SIP involves the case where a user in a different protocol domain generates a subscription that is destined for a user in a SIP domain.

Note that the process defined below requires that the gateway store subscription state. This unfortunate result is due to the need to remember the Call-ID, CSeq, and Route headers for subscriptions from the SIP side, so that they can be inserted into the SIP NOTIFY generated when a CPIM notification arrives.

7.1 SIP to CPIM

SIP for presnce is converted to CPIM through a SIP to CPIM abstract

gateway service, depicted in Figure 1.

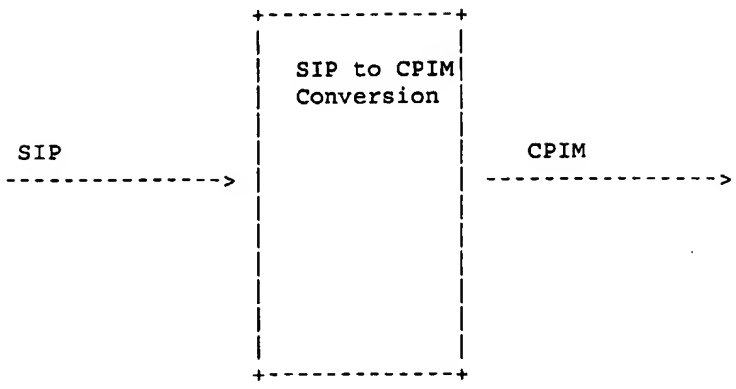


Figure 1: SIP to CPIM Conversion

The first step is that a SUBSCRIBE request is received at a gateway. The gateway generates a CPIM subscription request, with its parameters filled in as follows:

- The watcher identity in the CPIM message is copied from the From field of the SUBSCRIBE. If the From field contains a SIP URL, it is converted to an equivalent pres URL by dropping all SIP URL parameters, and changing the scheme to pres.

  This conversion may not work - what if the SIP URL has no user name. Plus, converting from a URL back to a URN in this fashion may not do it correctly.

- The target identity in the CPIM message is copied from the Request-URI field of the SUBSCRIBE. This may need to be converted to a pres URL as well.

- The duration parameter in the CPIM message is copied from the Expires header in the SUBSCRIBE. If the Expires header specifies an absolute time, it is converted to a delta-time by the gateway. If no Expires header is present, one hour is assumed.

- The transID parameter in the CPIM message is constructed by appending the Call-ID, the URI in the To field, the URI in the From field, the CSeq and the tag in the From field, and the request URI, and computing a hash of the resulting string. This hash is used as the transID. Note that the request URI is included in the hash. This is to differentiate forked requests within the SIP network that may arrive at the same gateway.

The CPIM service then responds with either a success or failure. In the case of success, the SIP to CPIM gateway service generates a 202 response to the SUBSCRIBE. It adds a tag to the To field in the response, which is the same as the transID field in the success response. The 202 response also contains a Contact header, which is the value of the target from the SUBSCRIBE request. It is important that the Contact header be set to the target, since that makes sure that subscription refreshes have the same value in the request URI as the original subscription. The duration value from the CPIM success response is placed into the Expires header of the 202. The gateway stores the Call-ID and Route header set for this subscription.

If the CPIM service responds with a failure, the SIP to CPIM gateway generates a 603 response. It adds a tag to the To field in the response, which is the same as the transID field in the failure response.

When the CPIM system generates a notification request, the SIP to CPIM gateway creates a SIP NOTIFY request. The request is constructed using the standard RFC2543 [2] procedures for constructing a request within a call leg. This will result in the To field containing the watcher field from CPIM, and the From field containing the target field from the CPIM notification. The tag in the From field will contain the transID. The presence information is copied into the body of the notification. The Call-ID and Route headers are constructed from the subscription state stored in the gateway. If no notification has yet been generated for this subscription, an initial CSeq value

is selected and stored.

SUBSCRIBE refreshes are handled identically to initial subscriptions as above.

If a subscription is received with an Expires of zero, the SIP to CPIM gateway generates an unsubscribe message into the the CPIM system. The watcher parameter is copied from the From field of the SUBSCRIBE. The target parameter is copied from the Request URI field of the SUBSCRIBE. The transID is copied from the tag in the To field of the SUBSCRIBE request.

The response to an unsubscribe is either success or failure. In the case of success, a 202 response is constructed in the same fashion as above for a success response to a CPIM subscriber. All subscription state is removed. In the case of failure, a 603 response is constructed in the same fashion as above, and then subscription state is removed, if present.

7.2 CPIM to SIP

CPIM to SIP conversion occurs when a CPIM subscription request arrives on the CPIM side of the gateway. This scenario is shown in Figure 2.

The CPIM subscription request is converted into a SIP SUBSCRIBE request. To do that, the first step is to determine if the subscribe is for an existing subscription. That is done by taking the target in the CPIM subscription request, and matching it against targets for existing subscriptions. If there are none, it is a new subscription, otherwise, its a refresh.

If its a new subscription, the gateway generates a SIP SUBSCRIBE request in the following manner:

- The From field in the request is set to the watcher field in the CPIM subscription request
- The To field in the request is set to the target field in the CPIM subscription request
- The Expires header in the SUBSCRIBE request is set to the duration field in the CPIM subscription request
- The tag in the From field is set to the transID in the CPIM subscription request.

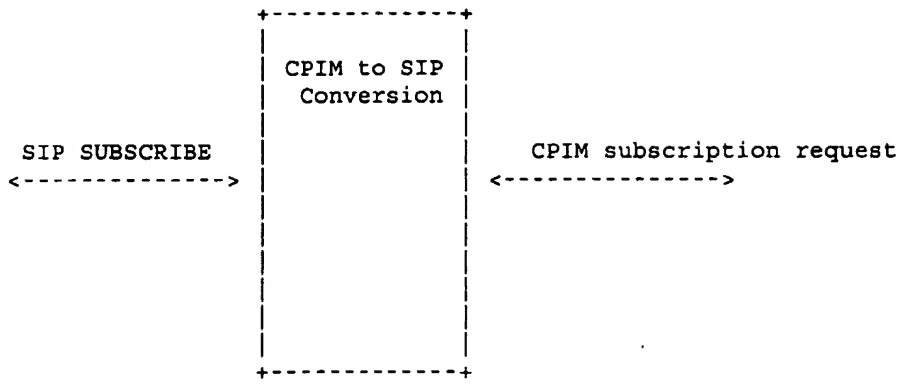


Figure 2: CPIM to SIP Conversion

This SUBSCRIBE message is then sent.

If the subscription is a refresh, a SUBSCRIBE request is generated in the same way. However, there will also be a tag in the To field, copied from the subscription state in the gateway, and a Route header, obtained from the subscription state in the gateway.

When a response to the SUBSCRIBE is received, a response is sent to the CPIM system. The duration parameter in this response is copied from the Expires header in the SUBSCRIBE response (a conversion from an absolute time to delta time may be needed). The transID in the response is copied from the tag in the From field of the response. If the response was 202, the status is set to indeterminate. If the response was any other 200 class response, the status is set to sucess. For any other final response, the status is set to failure.

If the response was a 200 class response, subscription state is

established. This state contains the tag from the To field in the SUBSCRIBE response, and the Route header set computed from the Record-Routes and Contact headers in the 200 class response. The subscription is indexed by the presentity identification (the To field of the SUBSCRIBE that was generated).

If an unsubscribe request is received from the CPIM side, the gateway checks if the subscription exists. If it does, a SUBSCRIBE is generated as described above. However, the Expires header is set to zero. If the subscription does not exist, the gateway generates a failure response and sends it to the CPIM system. When the response to the SUBSCRIBE request arrives, it is converted to a CPIM response as described above for the initial SUBSCRIBE response. In all cases, any subscription state in the gateway is destroyed.

When a NOTIFY is received from the SIP system, a CPIM notification request is sent. This notification is constructed as follows:

- The CPIM watcher is set to the URI in the To field of the NOTIFY.
- The CPIM target is set to the URI in the From field of the NOTIFY.
- The transID is computed using the same mechanism as for the SUBSCRIBE in Section 7.1
- The presence component of the notification is extracted from the body of the SIP NOTIFY request.

The gateway generates a 200 response to the SIP NOTIFY and sends it as well.

TODO: some call flow diagrams with the parameters

8 Firewall and NAT Traversal

It is anticipated that presence services will be used by clients and presentities that are connected to proxy servers on the other side of firewalls and NATs. Fortunately, since the SIP presence messages do not establish independent media streams, as INVITE does, firewall and NAT traversal is much simpler than described in [9] and [10].

Generally, data traverses NATs and firewalls when it is sent over TCP or TLS connections established by devices inside the firewall/NAT to devices outside of it. As a result, it is RECOMMENDED that SIP for presence entities maintain persistent TCP or TLS connections to their next hop peers. This includes connections opened to send a SUBSCRIBE,

NOTIFY, and most importantly, REGISTER. By keeping the latter connection open, it can be used by the SIP proxy to send messages from outside the firewall/NAT back to the client. It is also recommended that the client include a Contact cookie as described in [10] in their registration, so that the proxy can map the presentity URI to that connection.

Furthermore, entities on either side of a firewall or NAT should record-route in order to ensure that the initial connection established for the subscription is used for the notifications as well.

9 Security considerations

There are numerous security considerations for presence. Many are outlined above; this section considers them issue by issue.

9.1 Privacy

Privacy encompasses many aspects of a presence system:

- Subscribers may not want to reveal the fact that they have subscribed to certain users
- Users may not want to reveal that they have accepted subscriptions from certain users
- Notifications (and fetch results) may contain sensitive data which should not be revealed to anyone but the subscriber

Privacy is provided through a combination of hop by hop encryption and end to end encryption. The hop by hop mechanisms provide scalable privacy services, disable attacks involving traffic analysis, and hide all aspects of presence messages. However, they operate based on transitivity of trust, and they cause message content to be revealed to proxies. The end-to-end mechanisms do not require transitivity of trust, and reveal information only to the desired recipient. However, end-to-end encryption cannot hide all information, and is susceptible to traffic analysis. Strong end to end authentication and encryption also requires that both participants have public keys, which is not generally the case. Thus, both mechanisms combined are needed for complete privacy services.

SIP allows any hop by hop encryption scheme. It is RECOMMENDED that between network servers (proxies to proxies, proxies to redirect servers), transport mode ESP [11] is used to encrypt the entire message. Between a UAC and its local proxy, TLS [12] is RECOMMENDED. Similarly, TLS SHOULD be used between a presence server and the PUA.

The presence server can determine whether TLS is supported by the receiving client based on the transport parameter in the Contact header of its registration. If that registration contains the token "tls" as transport, it implies that the PUA supports TLS.

Furthermore, we allow for the Contact header in the SUBSCRIBE request to contain TLS as a transport. The Contact header is used to route subsequent messages between a pair of entities. It defines the address and transport used to communicate with the user agent. Even though TLS might be used between the subscriber and its local proxy, placing this parameter in the Contact header means that TLS can also be used end to end for generation of notifications after the initial SUBSCRIBE message has been successfully routed. This would provide end to end privacy and authentication services with low proxy overheads.

SIP encryption MAY be used end to end for the transmission of both SUBSCRIBE and NOTIFY requests. SIP supports PGP based encryption, which does not require the establishment of a session key for encryption of messages within a given subscription (basically, a new session key is established for each message as part of the PGP encryption). Work has recently begun on the application of S/MIME [13] for SIP.

9.2 Message integrity and authenticity

It is important for the message recipient to ensure that the message contents are actually what was sent by the originator, and that the recipient of the message be able to determine who the originator really is. This applies to both requests and responses of SUBSCRIBE and NOTIFY. This is supported in SIP through end to end authentication and message integrity. SIP provides PGP based authentication and integrity (both challenge-response and public key signatures), and http basic and digest authentication. HTTP Basic is NOT RECOMMENDED.

9.3 Outbound authentication

When local proxies are used for transmission of outbound messages, proxy authentication is RECOMMENDED. This is useful to verify the identity of the originator, and prevent spoofing and spamming at the originating network.

9.4 Replay prevention

To prevent the replay of old subscriptions and notifications, all signed SUBSCRIBE and NOTIFY requests and responses MUST contain a Date header covered by the message signature. Any message with a date

older than several minutes in the past, or more than several minutes into the future, SHOULD be discarded.

Furthermore, all signed SUBSCRIBE and NOTIFY requests MUST contain a Call-ID and CSeq header covered by the message signature. A user agent or presence server MAY store a list of Call-ID values, and for each, the higest CSeq seen within that Call-ID. Any message that arrives for a Call-ID that exists, whose CSeq is lower than the highest seen so far, is discarded.

Finally, challenge-response authentication (http digest or PGP) MAY be used to prevent replay attacks.

9.5 Denial of service attacks

Denial of service attacks are a critical problem for an open, inter-domain, presence protocol. Here, we discuss several possible attacks, and the steps we have taken to prevent them.

9.5.1 Smurf attacks through false contacts

Unfortunately, presence is a good candidate for smurfing attacks because of its amplification properties. A single SUBSCRIBE message could generate a nearly unending stream of notifications, so long as a suitably dynamic source of presence data can be found. Thus, a simple way to launch an attack is to send subscriptions to a large number of users, and in the Contact header (which is where notifications are sent), place the address of the target.

The only reliable way to prevent these attacks is through authentication and authorization. End users will hopefully not accept subscriptions from random unrecognized users. Also, the presence client software could be programmed to warn the user when the Contact header in a SUBSCRIBE is from a domain which does not match that of the From field (which identifies the subscriber).

Also, note that as described in [3], if a NOTIFY is not acknowledged or was not wanted, the subscription that generated it is removed. This eliminates the amplification properties of providing false Contact addresses.

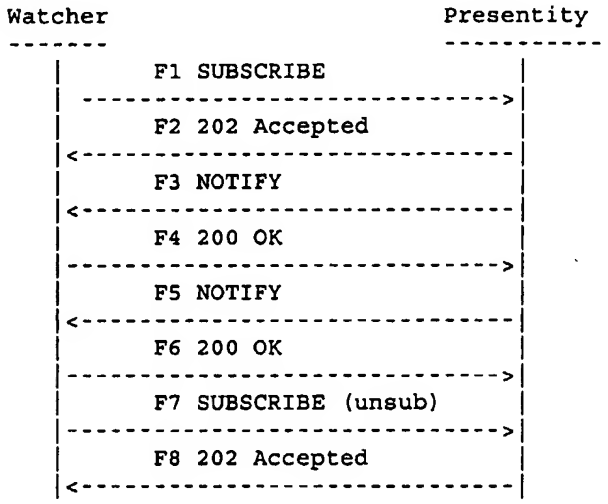
10 Example message flows

The following subsections exhibit example message flows, to further clarify behavior of the protocol.

10.1 Client to Client Subscription with Presentity State Changes

This call flow illustrates subscriptions and notifications that do not involve a presence server.

The watcher subscribes to the presentity, and the subscription is accepted, resulting in a 202 Accepted response. The presentity subsequently changes state (is on the phone), resulting in a new notification. The flow finishes with the watcher canceling the subscription.



Message Details

F1 SUBSCRIBE watcher -> presentity

```
SUBSCRIBE sip:presentity@pres.example.com SIP/2.0
Via: SIP/2.0/UDP  watcherhost.example.com:5060
From: User <pres:user@example.com>
To: Resource <pres:presentity@example.com>
Call-ID: 3248543@watcherhost.example.com
CSeq : 1 SUBSCRIBE
Expires: 600
Accept: application/xpidf+xml
Event: presence
Contact: sip:user@watcherhost.example.com
```

F2 202 Accepted presentity->watcher

```
SIP/2.0 202 Accepted
Via: SIP/2.0/UDP watcherhost.example.com:5060
From: User <pres:user@example.com>
To: Resource <pres:presentity@example.com>;tag=88a7s
Call-ID: 3248543@watcherhost.example.com
Cseq: 1 SUBSCRIBE
Event: presence
Expires: 600
Contact: sip:presentity@pres.example.com
```

F3 NOTIFY Presentity->watcher

```
NOTIFY sip:user@watcherhost.example.com SIP/2.0
Via: SIP/2.0/UDP pres.example.com:5060
From: Resource <pres:presentity@example.com>;tag=88a7s
To: User <pres:user@example.com>
Call-ID: 3248543@watcherhost.example.com
CSeq: 1 NOTIFY
Event: presence
Content-Type: application/xpidf+xml
Content-Length: 120

<?xml version="1.0"?>
<presence entityInfo="pres:presentity@example.com">
  <tuple destination="sip:presentity@example.com" status="open"/>
</presence>
```

F4 200 OK watcher->presentity

```
SIP/2.0 200 OK
Via: SIP/2.0/UDP pres.example.com:5060
From: Resource <pres:presentity@example.com>
To: User <pres:user@example.com>
Call-ID: 3248543@watcherhost.example.com
CSeq: 1 NOTIFY
```

F5 NOTIFY Presentity->watcher

```
NOTIFY sip:user@watcherhost.example.com SIP/2.0
Via: SIP/2.0/UDP pres.example.com:5060
From: Resource <pres:presentity@example.com>
To: User <pres:user@example.com>
Call-ID: 3248543@watcherhost.example.com
CSeq: 2 NOTIFY
Event: presence
Content-Type: application/xpidf+xml
Content-Length: 120

<?xml version="1.0"?>
<presence entityInfo="pres:presentity@example.com">
  <tuple destination="sip:presentity@example.com" status="closed"/>
</presence>
```

F6 200 OK watcher->presentity

```
SIP/2.0 200 OK
Via: SIP/2.0/UDP pres.example.com:5060
From: Resource <pres:presentity@example.com>
To: User <pres:user@example.com>
Call-ID: 3248543@watcherhost.example.com
CSeq: 2 NOTIFY
```

F7 SUBSCRIBE watcher -> presentity

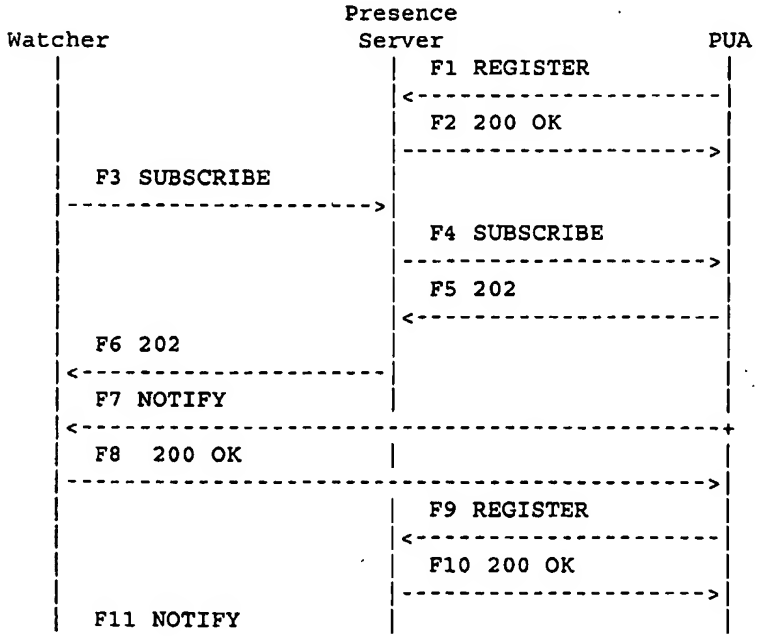
```
SUBSCRIBE sip:presentity@pres.example.com SIP/2.0
Via: SIP/2.0/UDP  watcherhost.example.com:5060
From: User <pres:user@example.com>
To: Resource <pres:presentity@example.com>
Call-ID: 3248543@watcherhost.example.com
Event: presence
CSeq : 2 SUBSCRIBE
Expires: 0
Accept: application/xpidf+xml
Contact: sip:user@watcherhost.example.com
```

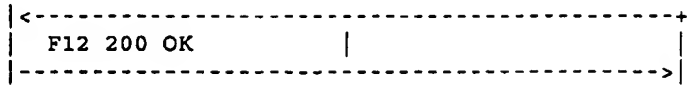
F8 202 Accepted presentity->watcher

```
SIP/2.0 202 Accepted
Via: SIP/2.0/UDP watcherhost.example.com:5060
From: User <pres:user@example.com>
To: Resource <pres:presentity@example.com>
Call-ID: 3248543@watcherhost.example.com
Event: presence
Cseq: 2 SUBSCRIBE
Expires:0
```

10.2 Presence Server with Client Notifications

This call flow shows the involvement of a presence server in the handling of subscriptions. In this scenario, the client has indicated that it will handle subscriptions and thus notifications. The message flow shows a change of presence state by the client and a cancellation of the subscription by the watcher.





Message Details

F1 REGISTER PUA->server

```
REGISTER sip:example.com SIP/2.0
Via: SIP/2.0/UDP pua.example.com:5060
To: <sip:resource@example.com>
From: <sip:resource@example.com>
Call-ID: 2818@pua.example.com
CSeq: 1 REGISTER
Contact: <sip:id@pua.example.com>;methods="MESSAGE"
         ;description="open"
Contact: <sip:id@pua.example.com>;methods="SUBSCRIBE"
Expires: 600
```

F2 200 OK server->PUA

```
SIP/2.0 200 OK
Via: SIP/2.0/UDP pua.example.com:5060
To: <sip:resource@example.com>
From: <sip:resource@example.com>
Call-ID: 2818@pua.example.com
CSeq: 1 REGISTER
Contact: <sip:id@pua.example.com>;methods="MESSAGE"
         ;description="open"
Contact: <sip:id@pua.example.com>;methods="SUBSCRIBE"
Expires: 600
```

F3 SUBSCRIBE watcher->server

```
SUBSCRIBE sip:resource@example.com SIP/2.0
Via: SIP/2.0/UDP  watcherhost.example.com:5060
From: User <pres:user@example.com>
To: Resource <pres:resource@example.com>
Call-ID: 32485@watcherhost.example.com
CSeq : 1 SUBSCRIBE
Expires: 600
Event: presence
Accept: application/xpidf+xml
Contact: sip:user@watcherhost.example.com
```

F4 SUBSCRIBE server->PUA

```
SUBSCRIBE sip:id@pua.example.com SIP/2.0
Via: SIP/2.0/UDP server.example.com:5060
Via: SIP/2.0/UDP watcherhost.example.com:5060
From: User <pres:user@example.com>
To: Resource <pres:resource@example.com>
Call-ID: 32485@watcherhost.example.com
CSeq : 1 SUBSCRIBE
Event: presence
Expires: 600
Accept: application/xpidf+xml
Contact: sip:user@watcherhost.example.com
```

F5 202 Accepted PUA->server

```
SIP/2.0 202 Accepted
Via: SIP/2.0/UDP server.example.com:5060
Via: SIP/2.0/UDP watcherhost.example.com:5060
From: User <pres:user@example.com>
To: Resource <pres:resource@example.com>;tag=ffd2
Call-ID: 32485@watcherhost.example.com
CSeq : 1 SUBSCRIBE
Event: presence
Expires: 600
```

F6 200 OK server->watcher

```
SIP/2.0 202 Accepted
Via: SIP/2.0/UDP watcherhost.example.com:5060
From: User <pres:user@example.com>
To: Resource <pres:resource@example.com>;tag=ffd2
Call-ID: 32485@watcherhost.example.com
CSeq : 1 SUBSCRIBE
Event: presence
Expires: 600
```

F7 NOTIFY PUA->watcher

```
NOTIFY sip:user@watcherhost.example.com SIP/2.0
Via: SIP/2.0/UDP pua.example.com:5060
To: User <pres:user@example.com>
From: Resource <pres:resource@example.com>;tag=ffd2
Call-ID: 32485@watcherhost.example.com
CSeq : 1 NOTIFY
Event: presence
Content-Type: application/xpidf+xml
Content-Length: 120

<?xml version="1.0"?>
<presence entityInfo="pres:resource@example.com">
  <tuple destination="im:resource@example.com" status="open"/>
</presence>
```

F8 200 OK watcher->PUA

```
SIP/2.0 200 OK
Via: SIP/2.0/UDP pua.example.com:5060
To: User <pres:user@example.com>
From: Resource <pres:resource@example.com>;tag=ffd2
Call-ID: 32485@watcherhost.example.com
CSeq : 1 NOTIFY
```

F9 REGISTER PUA->server

```
REGISTER sip:example.com SIP/2.0
Via: SIP/2.0/UDP pua.example.com:5060
To: <sip:resource@example.com>
From: <sip:resource@example.com>
Call-ID: 2818@pua.example.com
CSeq: 2 REGISTER
Contact: <sip:id@pua.example.com>;methods="MESSAGE"
         ;description="busy"
Contact: <sip:id@pua.example.com>;methods="SUBSCRIBE"
Expires: 600
```

F10 200 OK server->PUA

```
SIP/2.0 200 OK
Via: SIP/2.0/UDP pua.example.com:5060
To: <sip:resource@example.com>
From: <sip:resource@example.com>
Call-ID: 2818@pua.example.com
CSeq: 2 REGISTER
Contact: <sip:id@pua.example.com>;methods="MESSAGE"
         ;description="busy"
Contact: <sip:id@pua.example.com>;methods="SUBSCRIBE"
Expires: 600
```

F11 NOTIFY PUA->watcher

```
NOTIFY sip:user@watcherhost.example.com SIP/2.0
Via: SIP/2.0/UDP pua.example.com:5060
To: User <pres:user@example.com>
From: Resource <pres:resource@example.com>;tag=ffd2
Call-ID: 32485@watcherhost.example.com
CSeq : 2 NOTIFY
Event: presence
Content-Type: application/xpidf+xml
Content-Length: 120
```

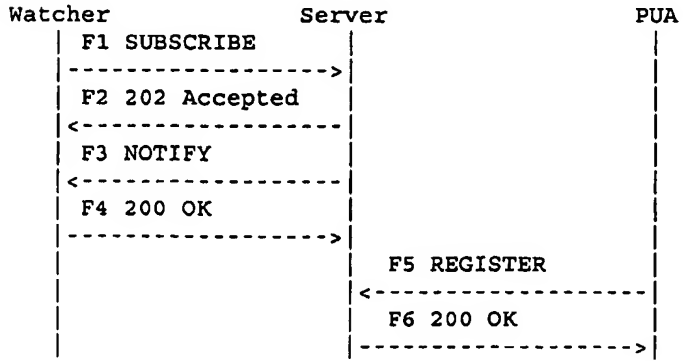
```
<?xml version="1.0"?>
<presence entityInfo="pres:resource@example.com">
  <tuple destination="im:resource@example.com" status="busy"/>
</presence>
```

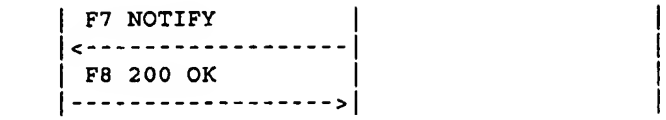
F12 200 OK watcher->PUA

```
SIP/2.0 200 OK
Via: SIP/2.0/UDP pua.example.com:5060
To: User <pres:user@example.com>
From: Resource <pres:resource@example.com>;tag=ffd2
Call-ID: 32485@watcherhost.example.com
CSeq : 2 NOTIFY
```

10.3 Presence Server Notifications

This message flow illustrates how the presence server can be the responsible for sending notifications for a presentity. The presence server will do this if the presentity has not sent a registration indicating an interest in handling subscriptions. This flow assumes that the watcher has previously been authorized to subscribe to this resource at the server.





Message Details

F1 SUBSCRIBE watcher->server

```
SUBSCRIBE sip:resource@example.com SIP/2.0
Via: SIP/2.0/UDP watcherhost.example.com:5060
To: <pres:resource@example.com>
From: <pres:user@example.com>
Call-ID: 2010@watcherhost.example.com
CSeq: 1 SUBSCRIBE
Event: presence
Accept: application/xpidf+xml
Contact: <sip:user@watcherhost.example.com>
Expires: 600
```

F2 202 OK server->watcher

```
SIP/2.0 202 Accepted
Via: SIP/2.0/UDP watcherhost.example.com:5060
To: <pres:resource@example.com>;tag=ffd2
From: <pres:user@example.com>
Call-ID: 2010@watcherhost.example.com
CSeq: 1 SUBSCRIBE
Event: presence
Expires: 600
Contact: sip:example.com
```

F3 NOTIFY server-> watcher

```
NOTIFY sip:user@watcherhost.example.com SIP/2.0
Via: SIP/2.0/UDP server.example.com:5060
```

```
From: <pres:resource@example.com>;tag=ffd2
To: <pres:user@example.com>
Call-ID: 2010@watcherhost.example.com
Event: presence
CSeq: 1 NOTIFY
Content-Type: application/xpidf+xml
Content-Length: 120

<?xml version="1.0"?>
<presence entityInfo="pres:resource@example.com">
  <tuple destination="im:resource@example.com" status="open"/>
</presence>
```

F4 200 OK

```
SIP/2.0 200 OK
Via: SIP/2.0/UDP server.example.com:5060
From: <pres:resource@example.com>;tag=ffd2
To: <pres:user@example.com>
Call-ID: 2010@watcherhost.example.com
CSeq: 1 NOTIFY
```

F5 REGISTER

```
REGISTER sip:example.com SIP/2.0
Via: SIP/2.0/UDP pua.example.com:5060
To: <sip:resource@example.com>
From: <sip:resource@example.com>
Call-ID: 110@pua.example.com
CSeq: 2 REGISTER
Contact: <sip:id@pua.example.com>;methods="MESSAGE"
         ;description="Away from keyboard"
Expires: 600
```

F6 200 OK

```
Via: SIP/2.0/UDP pua.example.com:5060
To: <sip:resource@example.com>
From: <sip:resource@example.com>
Call-ID: 110@pua.example.com
CSeq: 2 REGISTER
Contact: <sip:id@pua.example.com>;methods="MESSAGE"
          ; description="Away from keyboard"
          ; expires=600
```

F7 NOTIFY

```
NOTIFY sip:user@watcherhost.example.com SIP/2.0
Via: SIP/2.0/UDP server.example.com:5060
From: <pres:resource@example.com>;tag=ffd2
To: <pres:user@example.com>
Call-ID: 2010@watcherhost.example.com
CSeq: 2 NOTIFY
Event: presence
Content-Type: application/xpidf+xml
Content-Length: 120

<?xml version="1.0"?>
<presence entityInfo="pres:resource@example.com">
  <tuple destination="im:resource@example.com" status="Away from
keyboard"/>
</presence>
```

F8 200 OK

```
SIP/2.0 200 OK
Via: SIP/2.0/UDP server.example.com:5060
From: <sip:resource@example.com>;tag=ffd2
To: <pres:user@example.com>
Call-ID: 2010@watcherhost.example.com
CSeq: 2 NOTIFY
```

11 Open Issues

The following is the list of known open issues:

- o This draft recommends that the To and From field are populated with presence URLs rather than sip URLs. Is that reasonable? Will this lead to incompatibilities in proxies? Is there any issues with CPIM if the SIP URL format is used? This depends on what components of a message are signed in CPIM.
- o Rate limitations on NOTIFY: do we want that? How do we pick a value? 5 seconds is arbitrary.
- o Merging of presence data from multiple PA has been removed. Is that OK?
- o Placing IM URLs in the Contact header of a REGISTER: is that OK? What does it mean?
- o The process of migrating subscriptions from a presence server to PUA is not likely to work in the case where subscription refreshes use tags and Route headers. So, we have a choice. Either migration is disallowed, and we keep with leg oriented subscriptions, or migration is allowed, and there is no tags or Route's associated with subscriptions.
- o Converting SIP URLs back to pres URLs.
- o The SIP to CPIM and CPIM to SIP gateways are not stateless, because of the need to maintain Route, Call-ID, CSeq, and other parameters. Perhaps we can ask CPIM to define a token value which is sent in a CPIM request and returned in a CPIM response. Would that help?
- o Need to specify how to take Contacts from REGISTER and build a presence document. One obvious thing is that the contact addresses don't go in there directly; you probably want to put the address of record, otherwise calls might not go through the proxy.

12 Changes from -00

The document has been completely rewritten, to reflect the change from a sales pitch and educational document, to a more formal protocol specification. It has also been changed to align with the SIP event architecture and with CPIM. The specific protocol changes resulting from this rewrite are:

- The Event header must now be used in the SUBSCRIBE and NOTIFY requests.
- The SUBSCRIBE message can only have a single Contact header. -00 allowed for more than one.
- The From and To headers can contain presence URIs.
- The Request-URI can contain a presence URI.
- Subscriptions are responded to with a 202 if they are pending or accepted.
- Presence documents are not returned in the body of the SUBSCRIBE response. Rather, they are sent in a separate NOTIFY. This more cleanly separates subscription and notification, and is mandated by alignment with CPIM.
- Authentication is now mandatory at the PA. Authorization is now mandatory at the PA.
- Fake NOTIFY is sent for pending or rejected subscriptions.
- A rate limit on notifications was introduced.
- Merging of presence data has been removed.
- The subscriber rejects notifications received with tags that don't match those in the 202 response to the SUBSCRIBE. This means that only one PA will hold subscription state for a particular subscriber for a particular presentity.
- IM URLs allowed in Contacts in register
- CPIM mappings defined.
- Persistent connections recommended for firewall traversal.

Obtaining Authorization

When a subscription arrives at a PA, the subscription needs to be authorized by the presentity. In some cases, the presentity may have provided authorization ahead of time. However, in many cases, the subscriber is not pre-authorized. In that case, the PA needs to query the presentity for authorization.

In order to do this, we define an implicit subscription at the PA. This subscription is for a virtual presentity, which is the "set of

subscriptions for presentity X", and the subscriber to that virtual presentity is X itself. Whenever a subscription is received for X, the virtual presentity changes state to reflect the new subscription for X. This state changes for subscriptions that are approved and for ones that are pending. As a result of this, when a subscription arrives for which authorization is needed, the state of the virtual presentity changes to indicate a pending subscription. The entire state of the virtual presentity is then sent to the subscriber (the presentity itself). This way, the user behind that presentity can see that there are pending subscriptions. It can then use some non-SIP means to install policy in the server regarding this new user. This policy is then used to either accept or reject the subscription.

A call flow for this is shown in Figure 3.

In the case where the presentity is not online, the problem is also straightforward. When the user logs into their presence client, it can fetch the state of the virtual presentity for X, check for pending subscriptions, and for each of them, upload a new policy which indicates the appropriate action to take.

A data format to represent the state of these virtual presentities can be found in [14].

A Acknowledgements

We would like to thank the following people for their support and comments on this draft:

Rick Workman	Nortel
Adam Roach	Ericsson
Sean Olson	Ericsson
Billy Biggs	University of Waterloo
Stuart Barkley	UUNet
Mauricio Arango	SUN
Richard Shockey	Shockey Consulting LLC
Jorgen Bjorkner	Hotsip
Henry Sinnreich	MCI Worldcom
Ronald Akers	Motorola

B Authors Addresses

Jonathan Rosenberg

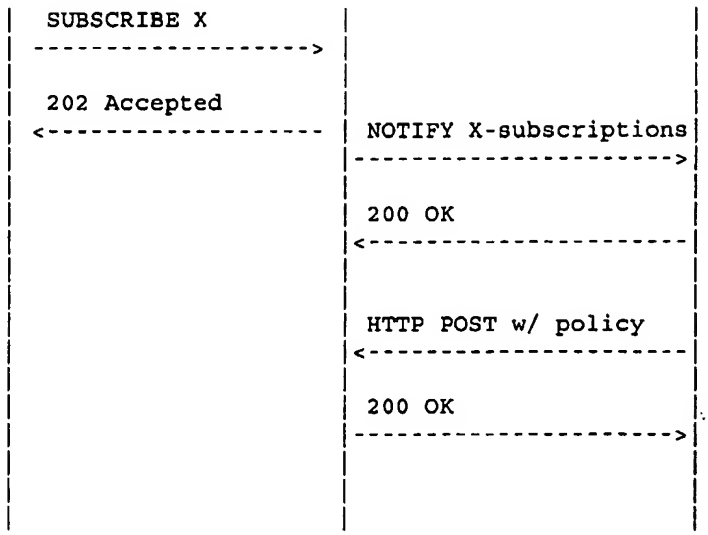


Figure 3: Sequence diagram for online authorization

dynamicsoft
72 Eagle Rock Avenue
First Floor
East Hanover, NJ 07936
email: jdrosen@dynamicsoft.com

Dean Willis
dynamicsoft
5100 Tennyson Parkway
Suite 1200
Plano, Texas 75024
email: dwillis@dynamicsoft.com

Robert Sparks
dynamicsoft
5100 Tennyson Parkway
Suite 1200
Plano, Texas 75024
email: rsparks@dynamicsoft.com

Ben Campbell
5100 Tennyson Parkway
Suite 1200
Plano, Texas 75024
email: bcampbell@dynamicsoft.com

Henning Schulzrinne
Columbia University
M/S 0401
1214 Amsterdam Ave.
New York, NY 10027-7003
email: schulzrinne@cs.columbia.edu

Jonathan Lennox
Columbia University
M/S 0401
1214 Amsterdam Ave.
New York, NY 10027-7003
email: lennox@cs.columbia.edu

Christian Huitema
Microsoft Corporation
One Microsoft Way
Redmond, WA 98052-6399
email: huitema@microsoft.com

Bernard Aboba
Microsoft Corporation

One Microsoft Way
Redmond, WA 98052-6399
email: bernarda@microsoft.com

David Gurle
Microsoft Corporation
One Microsoft Way
Redmond, WA 98052-6399
email: dgurle@microsoft.com

David Oran
Cisco Systems
170 West Tasman Dr.
San Jose, CA 95134
email: oran@cisco.com

C Bibliography

[1] M. Day, J. Rosenberg, and H. Sugano, "A model for presence and instant messaging," Request for Comments 2778, Internet Engineering Task Force, Feb. 2000.

[2] M. Handley, H. Schulzrinne, E. Schooler, and J. Rosenberg, "SIP: session initiation protocol," Request for Comments 2543, Internet Engineering Task Force, Mar. 1999.

[3] A. Roach, "Event notification in SIP," Internet Draft, Internet Engineering Task Force, Oct. 2000. Work in progress.

[4] D. Crocker et al. , "A common profile for instant messaging (CPIM)," Internet Draft, Internet Engineering Task Force, Nov. 2000. Work in progress.

[5] P. Calhoun, A. Rubens, H. Akhtar, and E. Guttman, "DIAMETER base protocol," Internet Draft, Internet Engineering Task Force, Sept. 2000. Work in progress.

[6] C. Rigney, S. Willens, A. Rubens, and W. Simpson, "Remote authentication dial in user service (RADIUS)," Request for Comments 2865, Internet Engineering Task Force, June 2000.

[7] J. Boyle, R. Cohen, D. Durham, S. Herzog, R. Rajan, and A. Sastry, "The COPS (common open policy service) protocol," Request for Comments 2748, Internet Engineering Task Force, Jan. 2000.

[8] H. Schulzrinne and J. Rosenberg, "SIP caller preferences and callee capabilities," Internet Draft, Internet Engineering Task Force, July 2000. Work in progress.

[9] J. Rosenberg, D. Drew, and H. Schulzrinne, "Getting SIP through firewalls and NATs," Internet Draft, Internet Engineering Task Force, Feb. 2000. Work in progress.

[10] J. Rosenberg and H. Schulzrinne, "SIP traversal through enterprise and residential NATs and firewalls," Internet Draft, Internet Engineering Task Force, Nov. 2000. Work in progress.

[11] S. Kent and R. Atkinson, "IP encapsulating security payload (ESP)," Request for Comments 2406, Internet Engineering Task Force, Nov. 1998.

[12] T. Dierks and C. Allen, "The TLS protocol version 1.0," Request for Comments 2246, Internet Engineering Task Force, Jan. 1999.

[13] B. Ramsdell and Ed, "S/MIME version 3 message specification," Request for Comments 2633, Internet Engineering Task Force, June 1999.

[14] J. Rosenberg et al. , "An XML based format for watcher information," Internet Draft, Internet Engineering Task Force, June 2000. Work in progress.

Table of Contents

3GPP TR 23.922 V1.0.0 (1999-10)

Technical Report

3rd Generation Partnership Project; Technical Specification Group Services and Systems Aspects; Architecture for an All IP network (3G TR 23.922 version 1.0.0)

Reference

DTR/TSGS-0223922U

Keywords

<UMTS, Release 2000, All IP Network>

3GPP

Postal address

3GPP support office address

650 Route des Lucioles - Sophia Antipolis
Valbonne - FRANCE
Tel.: +33 4 92 94 42 00 Fax: +33 4 93 65 47 16

Internet

http://www.3gpp.org

Contents

Foreword

This Technical Report has been produced by the 3GPP.

The contents of the present document are subject to continuing work within the TSG and may change following formal TSG approval. Should the TSG modify the contents of this TS, it will be re-released by the TSG with an identifying change of release date and an increase in version number as follows:

Version x.y.z

where:

x the first digit:

1 presented to TSG for information;

2 presented to TSG for approval;

3 Indicates TSG approved document under change control.

y the second digit is incremented for all changes of substance, i.e. technical enhancements, corrections, updates, etc.

z the third digit is incremented when editorial only changes have been incorporated in the specification;

1 Scope

This technical report will propose an architecture that provisions an all-IP architecture option for release 00. The purpose of the technical report is to

- identify key issues and affected ongoing 3GPP work that need to be resolved and
- propose a high level work plan for completion of an all IP release 00 UMTS standard

in order to provide this architectural option within Release 2000.

2 References

The following documents contain provisions which, through reference in this text, constitute provisions of the present document.

- References are either specific (identified by date of publication, edition number, version number, etc.) or non-specific.
- For a specific reference, subsequent revisions do not apply.
- For a non-specific reference, the latest version applies.
- A non-specific reference to an ETS shall also be taken to refer to later versions published as an EN with the same number.

[1] TS 22.101 version 3.6.0: Service Principles

[2] TS 23.121: Architectural Requirements for Release 99.

[3] TS 22.121: The Virtual Home Environment

[4] TS 23.002: Network architecture

[5] "Compressing TCP/IP Headers for Low-Speed Serial Links", IETF RFC 1144, V. Jacobson

3 Definitions and abbreviations

3.1 Definitions

For the purposes of the present document, the [following] terms and definitions [given in ... and the following] apply.

existing service: services supported in Release 99 and earlier releases for both GSM and UMTS.

All IP core network: core network of release 2000 that uses IP for transport of all user data and signalling

ERAN is defined as an evolved GSM BSS supporting EDGE modulation schemes on a 200kHz basis and real time packet services

3.2 Abbreviations

For the purposes of the present document, the following abbreviations apply:

<ACRONYM>	<Explanation>
2G	second generation
3G	third generation
AMR	Adaptive Multi Rate
AS	Application Server
BSC	Base Station Controller
BTS	Base Station
CAMEL	Customised Applications for Mobile Network Enhanced Logic
CAP	CAMEL Application Part
CC	Call Control
CCF	Call Control Function
CN	Core Network
CS	Circuit Switched
CSCF	Call State Control Function
CSE	CAMEL Service Environment
DN	Directory Number
DNS	Directory Name Server
EDGE	Enhanced Data for GSM
EGPRS	Enhanced GPRS
FFS	For Further Study
GGSN	Gateway GPRS Support Node
GMSC	Gateway MSC
GPRS	General Packet Radio Service
GSN	GPRS Support Node
GTP	GPRS Tunnelling Protocol
H-CSCF	Home CSCF
HN	Home Network
HSS	Home Subscriber Server
ICGW	Incoming Call Gateway
IN	Intelligent Network
INAP	IN Application Part
IP	Internet Protocol
ISDN	Integrated Services Digital Network
ISP	Internet Service Provider
ISUP	ISDN User Part
LAN	Local Area Network
LN	Logical Name
MAHO	Mobile Assisted Handover
MAP	Mobile Application Part

MCU	Media Control Unit
MExE	Mobile Execution Environment
MGCF	Media Gateway Control Function
MGW	Media Gateway
MM	Mobility Management
MO	Mobile Originated
MRF	Media Resource Function
MSC	Mobile Switching Centre
MT	Mobile Terminated/Terminal
NPA	Numbering Plan Area
O&M	Operations and Maintenance
ODB	Operator Determined Barring
OSA	Open Service Architecture
PCU	Packet Control Unit
PDP	Packet Data Protocol
PDU	Packet Data Unit
PS	Packet Switched
PSTN	Public Switched Telephony Network
QoS	Quality of Service
R00	Release 2000
R99	Release 1999
RA	Routing Area
RAN	Radio Access Network
RLC/MAC	Radio Link Control/Media Access Control
RNC	Radio Network Controller
R-SGW	Roaming Signalling Gateway
RTP	Real Time Protocol
SAT	SIM Application Toolkit
SCF	Service Control Function (IN) and Service Capability Features (VHE/OSA)
SCP	Service Control Point
S-CSCF	Serving CSCF
SGSN	Serving GPRS Support Node
SIP	Session Initiated Protocol
SLA	Service Level Agreement
SN	Serving Network
SRNC	Serving Radio Network Controller
SSF	Service Switching Function
TCP	Transmission Control Protocol
TE	Terminal Equipment
T-SGW	Transport Signalling Gateway
UDP	User Datagram Protocol
UE	User Equipment
UMS	User Mobility Server
UTRAN	UMTS Terrestrial Radio Access Network
VHE	Virtual Home Environment
VLR	Visitor Location Register
VoIP	Voice over IP
VPN	Virtual Private Network
WAP	Wireless Application Protocol
WIN	Wireless IN (ANSI-41)

4 Requirements

In order for TSG-SA2 to conduct a study of the architecture issues relating to the introduction of an All IP architecture within UMTS, assumptions were made as to the requirements for this architecture. TSG-SA1 are invited to validate and to extend these requirements as part of the work package on requirements for Release 2000.

4.1 General

The aim of the all IP architecture is to allow operators to deploy IP technology to deliver 3rd Generation services, that is an architecture based on packet technologies and IP telephony for simultaneous real time and non real time services. This architecture should be based on an evolution from Release 99 specifications and should be compatible with IMT-2000, providing global terminal mobility (roaming) [1].

The IP network should provide wireless mobility access based on ERAN and UTRAN with a common core network, based an evolution of GPRS, for both. In this context, an E-GPRS radio access network is a 200kHz GSM based network supporting EDGE and evolved to support real time packet services. Although EDGE is not within the scope of 3GPP, there are requirements for the core network of the all IP architecture, to be common to both access technologies.

The characteristics of this network are

- Based on an evolution of GPRS
- Common network elements for multiple access types including UTRAN and ERAN
- Packet transport using IP protocols
- IP Client enabled terminals
- Support for voice, data, real time multimedia, and services with the same network elements.

The report also covers the support of CS services on IP technologies.

The benefits of this approach include

- Ability to offer seamless services, through the use of IP, regardless of means of access (e.g. common features used by subscribers whether accessing via conventional land telephony, cable, wireless, HIPERLAN 2 etc.)
- Synergy with generic IP developments and reduced cost of service
- Efficient solution for simultaneous multi-media services including voice, data, and advanced real time services.
- Higher level of control of services
- Integrated, and cost reduced OA&M through IP
- Take advantage of Internet applications by supporting terminals which are IP clients.
- Cost reduction through packet transport

4.1.1 General Requirements

1. The overall aim of the all IP network is to support similar services to GSM release '99 and new innovative services. Where appropriate these services should inter-work with exisiting GSM services.
2. In addition it should also possible to support existing (R99 and before) services/capabilities (speech, data, multimedia, SMS, supplementary services, VHE,...)in a manner that is transparent to the users of these services [1]. That is, the network needs to provide the service capabilities required in such a way as to support interworking of these services between the R00 all IP network option and the other family networks two domain architecture option (GSM pre Release 99, UMTS release 99).

3. The standard shall enable the all IP core network to support release 99 CS terminals. This shall be standardised in such a way as to allow operators to decide whether or not they wish to support Release 99 CS only terminals.

4. The support of existing services shall not preclude the extension of service capabilities possible through the use of an all IP architecture.

5. When the all IP networks are deployed, there will be services and databases provided for existing networks which are non-IP based e.g. local number portability, free phone numbers, specialised corporate services. The all IP architecture will need to be able to access these services.

6. R'00 all IP core network shall allow implementations having a CS and a PS domain, that are separated like both these domains in the R'99 architecture. This implementation allows the two domains to evolve independently, e.g. to combine an all IP R'00 PS domain with a STM based R'99 CS domain. Furthermore it shall be possible to implement a CS domain that uses all IP based architecture and in distinct service areas of the same network a CS domain based on ATM/STM. This allows a smooth migration to an all IP based core network.

   The R'00 all IP architecture shall support that all services share bearer level transport and bearer control.

   R'00 architecture shall allow an operator to migrate a R'99 network into a R'00 network, without need for change of transport network technology, node numbering scheme etc. R'00 networks shall also allow connection of R'99 UTRAN over Iucs, to provide the operator with flexibility in the network implementation.

   Note: In the general R'00 architecture other transport technology than IP shall be possible.

4.2 Service Capabilities

4.2.1 General

The following general service capabilities are identified:

1. Legal interception has to be possible in the R00 All-IP network.

2. Emergency calls shall be supported in the R00 All-IP network.

4.2.2 Basic requirements for the Service and Application Platforms

List of requirements on the "Application and Service" block:

1. Service capabilities are to be made available to the Applications through Service capability features;
2. Service capability features are provided by one or more service capabilities (possibly directly provided by network functions, e.g. HSS, CSCF etc.), as illustrated in Figure 4-1;

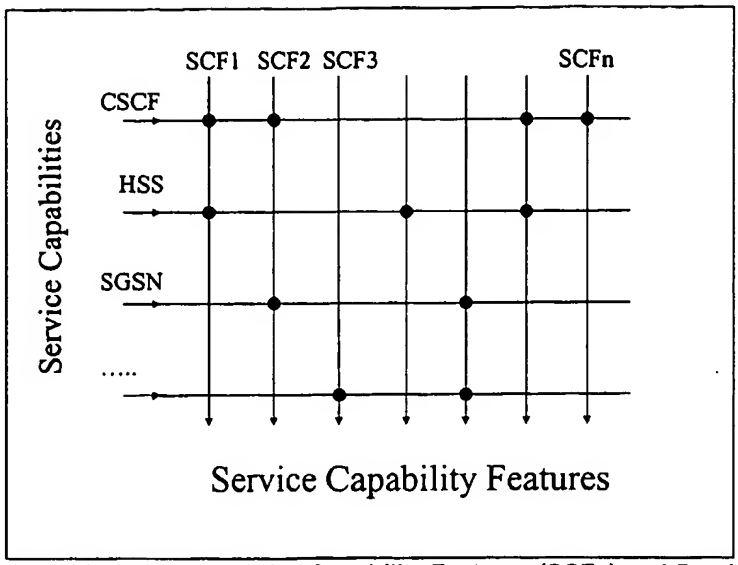


Figure 4-1: Relationship between Service Capability Features (SCFs) and Service Capabilities.

3. Incremental introduction of service capability features by the network operator has to be possible, for operator specific service capabilities;
4. A standard interface, called the application interface, has to be introduced for access to Service capability features by Applications. This interface has to provide a controlled, secure and accountable relationship between Applications and Service capability features;
5. The Application interface must provide access to the Service capability features based on user subscription profile;
6. Applications can be located either on servers and/or on (mobile) terminals;
7. Applications can only access the service capabilities through the service capability features;
8. Applications can utilize both capabilities provided by the Mobile Network functions, and functions as provided by IT systems, through the service capability features.

4.3 Numbering Schemes

The standards shall allow mobile terminated communications to be routed to the user's terminal on the basis of a single identifier e.g. MSISDN. This does not preclude multiple addresses being used for different services and capabilities (e.g. data, Fax, SMS). The network will route the call to the terminal over the available resources, dependent upon, for example, terminal capability, traffic loading and coverage. Networks migrating to an all IP architecture will require the ability to route based on a single identifier to maintain service transparency.

4.4 R99 Terminals

See section 4.1.1 above: The following requirements for the support of R99 terminals is an operator specific option.

1. The standards shall enable the All-IP core network to support UMTS R99 terminals.

2. Speech services including emergency calls shall be provided in All-IP networks to any UMTS R99 terminal supporting these services.

3. To ensure roaming of non VoIP capable UMTS R99 terminals, speech services including emergency calls shall be possible based on CS capabilities of these terminals.

4.5 Radio aspects

1. The radio resource usage should be optimised within the architecture for both service and signalling support.
2. Separation of the radio related and radio un-related functionalties between the core network and the radio access network
3. Separation of the user plane and control plane protocol stacks in the radio access networks
4. Fast uplink access and fast resource assignment procedures in both uplink and downlink for multiplexing different types of traffic on the same air link.
5. Optimization of end-to-end IP transport for certain class of real time applications (e.g. header compression, header stripping)
6. Network controlled handover procedure with short interruption to support real time applications (see handover requirement, section 4.9.)
7. Protocol stacks in the access network to support a range of services with different QoS requirements
8. Inter-working/interoperability of the QoS mechanism developed for the radio access network and the QoS mechanism used in the packet core network.
9. Bearer differentiation capability at the access network for multiplexing different types of traffic on the air to achieve maximum spectrum utilization
10. Optimal coding and interleaving for some applications such as voice.
11. Support of multiple data flows with different QoS per IP address (as defined in QoS framework in release 99)
12. Spectrum efficiency shall be maximized (e.g. statistical multiplexing).
13. The ERAN shall support GPRS and EGPRS services for pre-Release 2000 terminals.

4.6 Interworking

1. The All-IP core network shall support interworking to external IP and non-IP networks (e.g. circuit-switched networks (PSTN, ISDN, GSM PLMN, UMTS PLMN,...).

4.7 Mobility management

1. The All-IP core network shall provide streamlining and CN operated hand-over procedures for UMTS.

4.8 Roaming

1. The standard shall enable the All-IP core network to support roaming with 2G GSM/GPRS networks and R99 UMTS networks.

4.9 Handover

The support of handover between release 98, release 99 and release 2000 network technologies is essential in maintaining adequate network coverage. Table 4-1 illustrates the necessary handover scenarios and the status of development of mechanisms.

Table 4-1: Handover requirements for UMTS All IP network

Between	2G-GSM cs	2G-GPRS	UMTS cs (R99)	UMTS ps (R99)	UMTS IP (PS services)	UMTS IP (CS services)
UMTS IP (PS services)	Req R00*	Req R00	Req R00*	Req R00	OK	Not Required
UMTS IP (CS services)	Req R00	Not Req	Req R00	Not Req	Not Req	OK

Key:

OK Same technology

Req R00 Required for release 2000

* The implications of the requirement for PS to/from CS handover in R00 are the subject of much debate. Alternatives for either the support of handover or to provide service coverage need to be investigated.

4.9.1 Handover Categories

1 Intra network handover

Handover inside one all IP network

1a Intra RAN handover

1b Inter RAN handover

2 Inter network handover

Handover between two all IP networks

3 Inter-system handover

Handover between an all IP network and other systems

4.9.2 General Definition

Reselection and handover are two methods of supporting mobility during an active session. Reselection is the process whereby the mobile station autonomously determines which cell the mobile will receive services on. Handover is the process whereby the network determines which cell the mobile will receive services on.

4.9.3 General Requirements

For real-time services handover procedure shall be used. The network shall control the handover procedure.

Handover shall be the selected method of mobility if one or more active sessions have requested handover in a multi session call with different QoS requirements

Performance requirement on speech interruption (i.e. "mute" period) shall be equivalent to or better than than GSM or ANSI-136 handover for telephony services. TBD for other services offered by an all IP Network.

Maintain maximum packet loss limit (i.e. less than TBD) and maximum delay limit (i.e., less than TBD) during handover.

Non-real-time services shall use either handover or cell reselection depending on the QoS parameters in combination with network parameters.

The Subscribed QoS level should be maintained across the handover boundary. However QoS negotiation (if neceassary) should be possible before, during and after the handover (The application may reject the offered QoS)

Handover procedure shall utilize radio resources efficiently.

Handover shall not compromise the security of: the network providing the new radio resources; the (possibly different) network providing the original radio resources; and the terminal UE.

There shall be efficient handling of multiple bearers, e.g. if voice and email transfer is going on simultaneously.

Essential IP/UDP/RTP header information (for inter and intra all IP network handover), as seen by an IP end point, shall be preserved across handover boundary. The required essential header information depends on the bearer.

4.9.4 MS Requirements

The mobile station shall be capable of supporting both reselection and handover.

The mobile station shall aid the RAN in the handover decision by supplying RF environmental information (e.g. received signal strength from serving cell and neighbour cells).

4.9.5 RAN Requirements

Handover decisions shall be based in the RAN.

Maintain the RAN QoS parameters, associated with the mobile station, across a handover boundary. Note, RAN QoS parameters for a mobile station are based upon the negotiated set of QoS parameters.

Facilitate admission control to optimize radio resources.

Select a handover target based on criteria such as RF environmental information, radio resources of the neighboring cells, QoS requirements of the session, etc.

4.10 Call Control and Roaming

The following requirements need consideration for call control and roaming support in an all IP based network.

1 Routing of signalling and transport needs to be optimised, for the purposes of call control and roaming between networks.

2 Whenever possible, tromboning of the user's voice or data communication session back to their home environment should not be used to provide the user with services when roaming outside their home network.

3 The Release 2000 all IP network must comply with the mandated requirements for Emergency Services.

4 The Release 2000 all IP network must comply with the mandated requirements for Number Portability.

5 The Release 2000 all IP network must support multiparty voice and data communications sessions including the capability for the user or service logic to dynamically add or delete users from an active communications session.

6 The Release 2000 all IP network must be able to accept and re-route incoming voice or data communication requests that are addressed to the user's directory number during periods of realignment of the national numbering plans (e.g., NPA splits in North America).

7 Transcoding of the traffic (voice, data, video) should be minimised. For example,, if the terminal equipment of the called and calling party have the same vocoder, no transcoding of the voice traffic, within the network, would occur.

8 The Release 2000 all IP network must provide connection to the services of the legacy 2G and release 99 networks.

9 The Release 2000 all IP network shall support VHE for roamers.

10 A minimum set of services for roamers shall be provided within the serving network. This minimal set of user services is still being defined. However, the following is anticipated to be in this minimal set of user services:

a Speech call and data session origination.

b Speech call and data session termination.

c Call Waiting for voice calls in the case of monomedia session

d Call Forwarding services for voice calls.

e Calling party identification information

f SMS

11 In the event that the Release 2000 all IP operator does not have a legacy network in the market that a Release 2000 all IP network is being deployed into and the Release 2000 all IP operator does not have any business relationships with the operators of the legacy networks. Consequently, the design of the Release 2000 all IP network can not assume that the requirements for mandated services or operator-specific services can be satisfied by forwarding the Release 2000 all IP call to the legacy network. The following are examples of Operator Services that may need to be handled directly by the Release 2000 all IP networks:

a Directory Assistance

b Third party billing

c Collect calls

d Calling card calls

12 When a Release 2000 all IP user roams from a Release 2000 all IP network to another Release 2000 all IP network and gets access to both transport services (e.g. GPRS) and application level services (e.g. multimedia calls), services may be provided by a CSCF in the serving network or by a CSCF in the home network. The serving network shall contain the information to contact the user's home network for the user's profile information. The CSCF of the serving network shall have access to the necessary information for the invocation and control of the user's advanced/ supplementary services at the user's home network.

13 The user shall be able to gain access to their ISPs or corporate LAN application level services..

14 Both dynamic and dedicated IP addresses shall be supported.

15 Release 2000 all IP networks will be capable of providing VPN functionality. VPN refers to both the GSM VPN and Intranets. The VPN features supported require further study and analysis.

16 When the UE uses the service of a CSCF/MSC server, the CSCF/ MSC Server needs to authenticate the UE. The way this is handled when the UE uses the services of a CSCF in the visited network requires further study and analysis.

4.11 Security

The general principle for security for the all IP network implementation is to reuse the same mechanisms developed for 3GPP Release 99 wherever possible.

5 Architecture for an all IP PLMN

5.1 Reference Architecture

The reference architecture provides two options:

Option 1: has been developed with the goal of allowing operators to deploy an all IP based architecture to deliver 3rd Generation wireless/mobile services. This architecture is based on packet technologies and IP telephony for simultaneous real time and non real time services.

Option 2:. One purpose of option 2 is also to allow support of release 99 CS domain terminals. In addition option 2 also supports the IP based services of option 1.

5.1.1 Reference Architecture – Option 1

As described earlier in the Requirements section 4.1, the architecture shown in Figure 5-1 has been developed with the goal of allowing operators to deploy an all IP based architecture to deliver 3rd Generation wireless/mobile services. This architecture is based on packet technologies and IP telephony for simultaneous real time and non real time services.

The architecture shown and the components of which are described in subsequent sections allow for flexible and scaleable mechanisms to support global roaming and interoperability with external networks such as PLMN, 2G Legacy networks, PDNs and other multimedia VOIP networks.

The end-to-end architecture consists of the following key segments:

a) Radio Network

b) The GPRS network

c) The Call Control

d) Gateways to external networks

e) The Service architecture

The Radio network part consists of the equipment associated with the mobile user, the Radio Airlink and the Radio Access Network. The RAN supports both the UTRAN and the EDGE technologies.

Section 4 indicates that the intent of the core network part of the all IP architecture is that it should be designed to allow operators to use other access networks, for example ERAN and HIPERLAN 2. Within Figure 5-1, the ERAN is shown explicitly, where as the other access networks are represented by the bubble labelled "Alternative Access Network". For the purposes of this report,

> *the ERAN is defined as an evolved GSM BSS supporting EDGE modulation schemes on a 200kHz basis and real time packet services*

The support of alternative access networks, and the impact of the All IP architecture on the ERAN are outside the scope of this activity. However, to avoid the loss of information, the report does indicated where requirements are known to apply to these access networks.

Within this report, the reference point between the ERAN and the core network is designated as the Iu_ps'. That is, the reference point is Iu and the implementation is expected to be similar to that of the Iu_ps.

Note: the use of the reference label Iu_ps' is confusing. During the standardisation activity a more suitable label should be chosen.

The GPRS network part has the GSNs which provide the mobility management and the PDP context activation services to the mobile terminal as they do in the R99 GPRS PS domain network. The HLR functionality for the GPRS network is provided by the Home Subscriber Server, (HSS).

The Call Control part of the architecture is the most critical functionality. The CSCF, MGCF, R-SGW, MGW, T-SGW and the MRF comprise the Call Control and signalling functionality to deliver the real-time mobile/wireless services. The CSCF is similar to the H.323 GateKeeper or a SIP Server. The architecture has been intentionally kept generic and is not based on a specific call control mechanisms such as H.323 or SIP. Such a choice is for further study.

The user profiles are maintained in the HSS. The signalling to the multimedia IP network is interface solely via the CSCF while the bearer is interfaced directly with the GGSN. The MRF interfaces with all bearer components for bearer media and with the CSCF for signaling. The MRF provides for media mixing, multiplexing, other processing and generation functions.

The interconnectivity to external networks such as PLMN, other PDNs, other multimedia VOIP networks and 2G Legacy networks (GSM or TDMA) is supported by the GGSN, MGCF, MGW, R-SGW and T-SGW functional elements. Other PLMNs are interfaced for both bearer and signalling via their respective GPRS components. The CSCF is a new component which also participates in this signalling. The signalling to legacy mobile networks is interfaced

via the R-SGW, CSCF, MGCF, T-SGW and HSS, while the bearer is interfaced to and from the legacy PSTN network via an MGW. Legacy landline circuit switch signalling is interfaced via the CSCF, MGCF and T-SGW while the bearer is interfaced to and from the legacy PSTN network via an MGW.

The Service Architecture part of the network is currently depicted as an external entity and is described in detail in the section 10. Non-standard services are provided via interfaces to an application services layer. The HSS, the SGSN and the CSCF interface with the application and services bubble.

The details for each of the functional elements of the architecture are provided in subsequent subsections of this document.

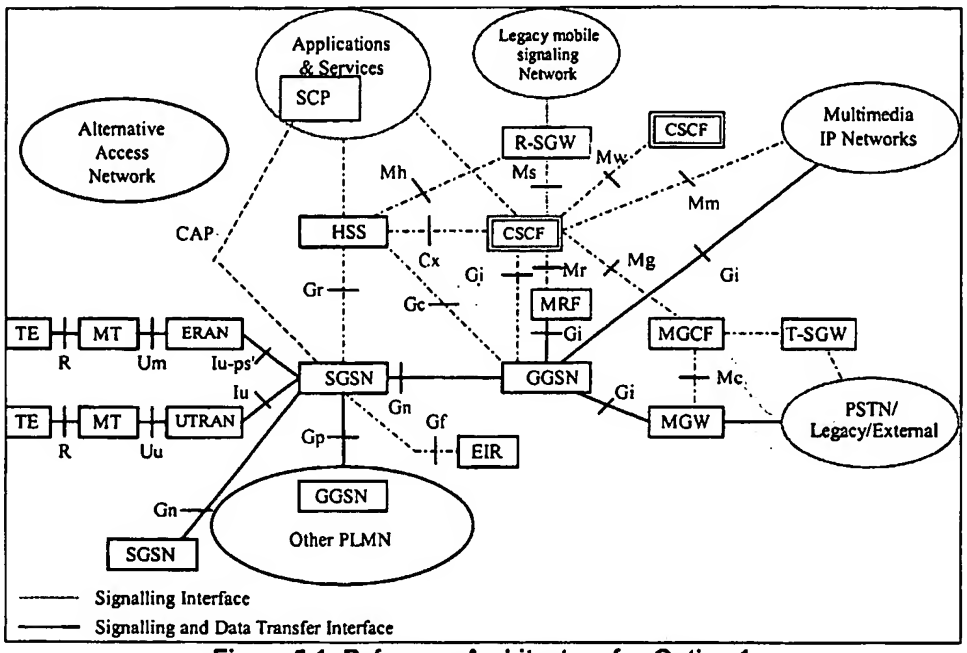


Figure 5-1: Reference Architecture for Option 1

The Gm interface between the UE and the CSCF consists of the user to network multimedia signalling. It is carried on radio, Iu, Gn and Gi interfaces.

The SGSN and GGSN are the same functional elements as defined in 23.002 [4] for R99 of UMTS.

Mobility management procedures for Release 2000 all IP network MSs in Release 2000 all IP networks are based on the Routing Area identifier (RAId). Mobility management procedures for CS capable MSs in Release 2000 all IP networks are based on the Routing Area identifier (RAId) and on Location Area Identifiers. In case of roaming between Release 2000 all IP networks to 2G and vice versa, a mechanism to convert the format of identities is needed (i.e. MS roaming from Release 2000 all IP network to 2G network has knowledge only of old RA but Location Area Identifier also needs to be given to the 2G-MSC). When roaming from 3G-R00 to 2G (without possibility of combined update), how the 2G-MSC can retrieve the IMSI from the all IP core network is an open issue.

MS identity (IMSI) is protected over the radio interface through the adoption of a single temporary identifier (P-TMSI) allocated by the SGSN during the MS registration procedure. P-TMSI can be reallocated at every following registration or routing area update. [relevance to roaming scenarios between Release 2000 all IP networks and legacy cellular networks]

The Access Network Nodes (GSN(s), RNC) are not aware the multimedia signaling protocol between the UE and the CSCF. They are even not aware that a given UE sends or does not send signalling to the CSCF.
Note1: this does not preclude that to optimize the radio, the RNC might support specific RAB for the individual flows of the multimedia user plane. These RAB are requested by the UE at PDP context activation.

Note2: the SGSN may have a role in the choice of the CSCF by the UE. This is FFS. But even if the SGSN participate in the CSCF address determination, the SGSN does not carry out the multimedia registration on behalf of the UE.

Different PDP contexts carry multimedia signalling and user flows due to different requirements on QoS for these PDP contexts. The Access Network Nodes (GSN(s), RNC) are nevertheless not aware whether a given PDP context carries multimedia signalling or not.

Working Architectural approach

1. The All IP Core network is engineered primarily to use a common technology (IP) to support all services including multimedia and voice services controlled by H.323/SIP or ISUP.
2. Network architecture is based upon IP packet technologies for simultaneous real-time and non-real-time services.
3. Network architecture is based upon an evolution of GPRS.
4. For support of R99 CS domain services the R99 CS domain CC mechanism may be reused. (NOTE: This does not prevent alternative mechanism such as H.323, SIP or evolved forms of R99 CS domain CC mechanisms being used by operators to deliver R99 CS domain services)
5. For the support of release 00 terminals are IP based, and the integration of services is obtained through IP.
6. Network architecture should support personal mobility and interoperability between mobile and fixed networks for both voice and data services.
7. Maintain or improve quality of service levels when compared to today's networks.
8. Maintain or improve network reliability when compared to today's networks.
9. All IP interfaces and associated network interfaces should be enhanced to support real-time multimedia services.
10. Network architecture will provide a separation of service control from call/connection control.
11. Network architecture will replace SS7 transport with IP.
12. Network architecture will be independent of network transport layers of Layer 1 (L1) and Layer 2 (L2).
13. Regardless of service type, ISUP based or IP based, IP transport shall be possible for all signalling and data transport.

5.1.2 Reference Architecture – Option 2

As described earlier in the Requirements section 4.1.1 item 3 and 6, the architecture shown in Figure 5-2 allows operators to migrate from a R'99 UMTS network into a R'00 All IP network. One purpose of option 2 is to allow support of release 99 CS terminals. Option 2 allows the two domains of R'99 to evolve independently.

As for option 1 the architecture of option 2 allows that all services supported by option 2 share bearer level transport and bearer control. Various underlaying transport mechanisms shall be allowed (e.g. RTP/UDP/IP, AAL2/ATM or STM).

Reference architecture option 2 includes the SGSN/GGSN/CSCF based services of reference architecture option 1. The definitions and working architectural approach described in chapter 5.1.1 therefore also applies to the SGSN/GGSN/CSCF part of option 2.

[Note: The requirement in section 4.3 for the support of routing on a single MSISDN or to allow operators to move subscribers from the CS services to the All IP service without changing MSISDN will be solved as part of R99.]

Two control elements, related with R'99 CS domain architecture, are added in option 2; the MSC server and GMSC server.

Option 2 benefits from the Iu architecture of R'99, having transport of user data separated from control, to allow UTRAN to access the core network via a MGW separated from the MSC server. Between UTRAN and MSC server the control part of Iu, RANAP, is used.

[Note: 1) Exact definition of ERAN *in relation to GSM/teleservice speech and ERAN's relations to MSC server/MGW needs to be defined. 2) The possibility of interfacing GSM/BSS to MGW and MSC server shall be further studied.].*

By allowing servers to terminate MAP and the user-network signalling (04.08+ CC+MM), requirements related to service and network migration of R'99 UMTS CS domain services and network migration can be fulfilled. The requirements that requires a network architecture according to option 2 are:

Section	Requirement Number
4.1.1	2 (to meet the timescales of R00)
4.1.1	3
4.1.1	6
4.4	all
4.8	1 (the need for option 2 to meet this requirement will depend upon the roaming solution chosen)

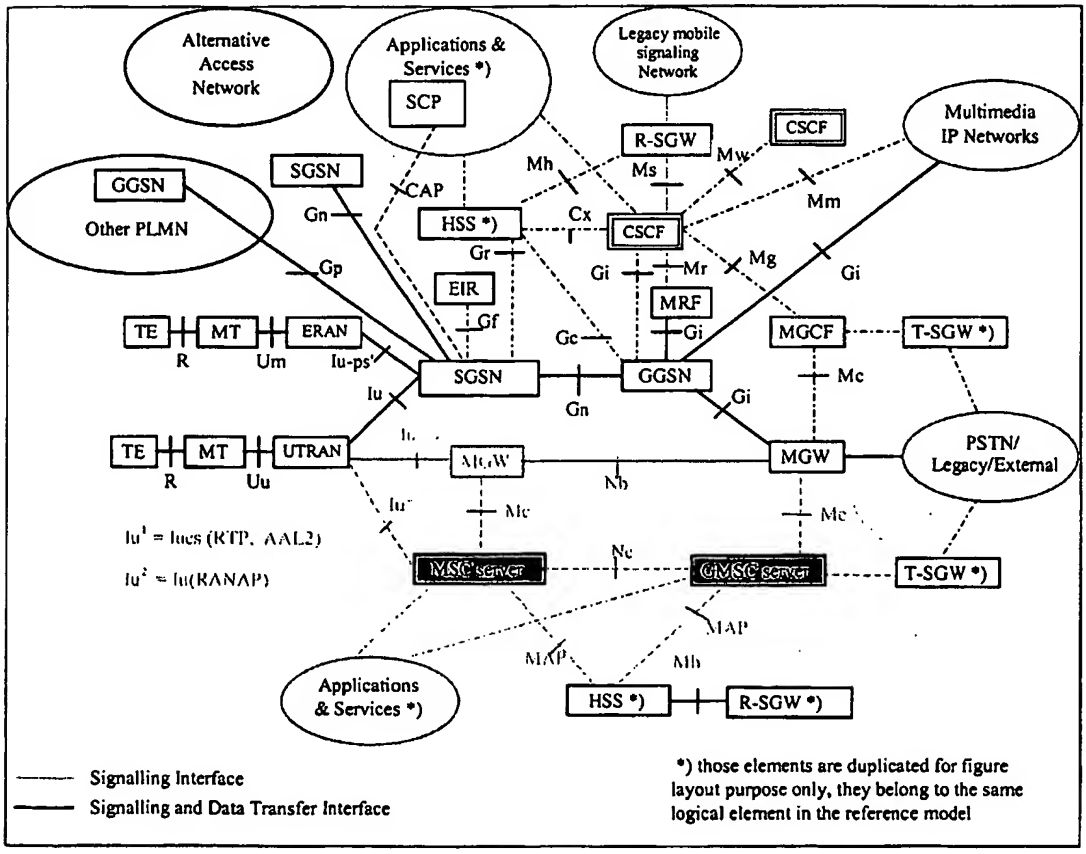


Figure 5-2: Reference Architecture for Option 2

Iu is the reference point between UTRAN and all IP core network. Between UTRAN and SGSN Iu is IP based. Between UTRAN and MGW - Iucs (RTP, AAL2) - Iu may be based on different transport technologies.

MAP is operated between HSS and MSC server and GMSC server respectively.

5.1.3 What is the Border of the Network

Open Issue: What is the Border of the Network

In the case that the GGSN is seen as the border of the network towards the IP network or the GGSN+MGW is seen as the network border towards the PSTN/Legacy network, then the following issue need to be clarified: how to determine the MGW and how to ensure the most optimal routing towards this MGW. On call setup, the CSCF needs to determine the appropriate MGW for the call. For example, it needs to determine if the call is to the PSTN (and in this case towards which PSTN network) or to a Voice Over IP network. This determination can only be ensured when call set-up signalling has been analysed by the CSCF and possibly by the SCF. This analysis may change the called party number (for instance modify the called party address from the address corresponding to an IP terminal to an address corresponding to a foreign PSTN terminal). It is only at that point that the best MGW can be determined. This determination cannot be done before by the SGSN. The MGW address is sent back (through H.225 signalling) to the UE. The UE can then activate a PDP context (for the support of user plane traffic) towards the appropriate network (i.e. the network that best allows to reach the MGW to be used by the call).

Note [FFS]. The issue of optimal routing when an MRF is added cannot be determined until there is agreement on what is the border of the network.

5.2 New Functional Elements

5.2.1 Call State Control Function (CSCF)

In the following section, CSCF has been divided into several logical components.

Currently, these logical components are internal to the CSCF. The need for external components to be able to address directly one of the logical components of the CSCF is for FFS.

Every CSCF acting as a Serving CSCF (see section 9) has a CCF function.

ICGW (Incoming call gateway)

- Acts as a first entry point and performs routing of incoming calls,
- Incoming call service triggering(e.g. call screening/call forwarding unconditional) may need to reside for optimisation purposes,
- Query Address Handling (implies administrative dependency with other entities)
- Communicates with HSS

CCF (Call Control Function)

- Call set-up/termination and state/event management
- Interact with MRF in order to support multi-party and other services
- Reports call events for billing, auditing, intercept or other purpose
- Receives and process application level registration
- Query Address Handling (implies administrative dependency)
- May provide service trigger mechanisms (service capabilities features) towards Application & services network (VHE/OSA)
- May invoke location based services relevant to the serving network

- May check whether the requested outgoing communication is allowed given the current subscription.

SPD (Serving Profile Database)

- Interacts with HSS in the home domain to receive profile information for the R00 all-IP network user and may store them depending on the SLA with the home domain
- Notifies the home domain of initial user's access (includes e.g. CSCF signalling transport address, user ID etc. needs further study)
- May cache access related information (e.g. terminal IP address(es) where the user may be reached etc.)

AH (Address Handling)

- analysis, translation, modification if required, address portability, mapping of alias addresses
- May do temporary address handling for inter-network routing.

Other functions such as admission control, multiple session knowledge within one CSCF, multiple CSCFs serving one terminal for multiple services, role of multiple CSCFs serving a network etc. need further investigation.

Interfaces need to be further studied and defined.

5.2.2 Home Subscriber Server (HSS)

The **Home Subscriber Server** (HSS) is the master database for a given user. It is responsible for keeping a master list of features and services (either directly or via servers) associated with a user, and for tracking of location of and means of access for its users. It provides user profile information, either directly or via servers. It is a superset of the Home Location Register (HLR) functionality., for example as defined in GSM MAP, but differs in that it needs to also communicate via new IP based interfaces. The HSS shall support a subscription profile which identifies for a given user for example:

- user identities
- subscribed services and profiles
- service specific information
- mobility management information
- authorization information

Like the HLR, the HSS contains or has access to the authentication centers/servers (e.g. AUC, AAA).

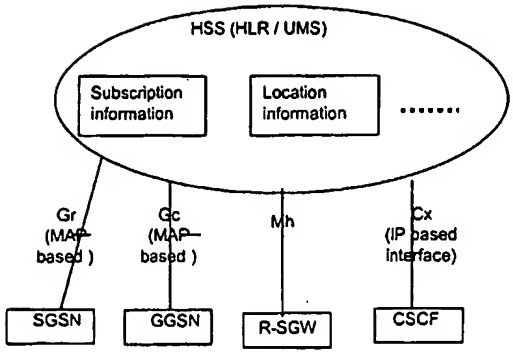


Figure 5-3: Example of a Generic HSS structure and basic interfaces

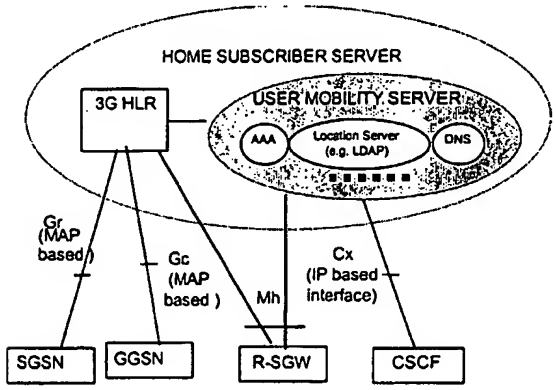


Figure 5-4: Example of HSS structure with UMS Specific Functionality

The HSS may consist of the following elements as shown in the Figure 3:

1) User Mobility Server (UMS): it stores the Release 2000 all IP network Service Profile(see section 9.1) and stores Service Mobility or Serving CSCF related information for the users. UMS might also generate, store and/or manage security data and policies (e.g. IETF features). UMS should provide logical name to transport address translation in order to provide answer to DNS queries. UMS role and functional decomposition are for further study.

2) 3G HLR: A GPRS HLR enhanced to support Release 2000 all IP networks GPRS specific information.

Gr and Gc use MAP which may be implemented using MAP transported over IP, however the issue of roaming to a network that supports MAP over SS7 needs to be considered. The Cx interface requires further study: it may be implemented via IETF protocols such as DNS or via MAP procedures.

Following functionality may need to be supported in the HSS and are for further study:

- it stores the R00 all- IP network Service Profile and stores location information for the users.
- it may also generate, store and/or manage security data and policies (e.g. IETF features).
- may need to provide logical name to transport address translation.
- The HSS interacts with the R-SGW to communicate with VLRs and other Mobility managers which do not use IP. HSS interfaces with CSCF via Cx which is for further study.
- Other R00 all-IP based network functions such as AAA, DNS etc. and their interactions with HSS is for further study.
- Interface(s) between UMS and 3G HLR is for further study.

note: If the user profile is split across different databases then there should either be no duplication of information elements or the consistency of the data should be maintained.

5.2.3 Transport Signalling Gateway Function (T-SWG)

This is component in the R00 all-IP network is PSTN/PLMN termination point for a defined network. The functionality defined within T-SGW should be consistent with existing/ongoing industry protocols/interfaces that will satisfy the requirements.

- Maps call related signalling from/to PSTN/PLMN on an IP bearer and sends it to/from the MGCF.
- Needs to provide PSTN/PLMN <-> IP transport level address mapping.

Interfaces need to be further studied and defined.

5.2.4 Roaming Signalling Gateway Function (R-SGW)

The role of the R-SGW described in the following bullets is related only to roaming to/from 2G/R99 CS and GPRS domain to/from R00 CS and GPRS domain and is not involving the multimedia/VoIP domain.

- In order to ensure proper roaming, the R-SGW performs the signaling conversion at transport level (conversion: Sigtran SCTP/IP versus SS7 MTP) between the legacy SS7 based transport of signaling and the IP based transport of signaling. The R-SGW does not interpret the MAP / CAP messages but may have to interpret the underlying SCCP layer to ensure proper routing of the signaling.

- (For the support of 2G / R99 CS terminals): The services of the R_SGW are used to ensure transport interworking between the SS7 and the IP transport of MAP_E and MAP_G signalling interfaces with a 2G / R99 MSC/VLR

For the Multimedia/VoIP domain, MAP interworking at the R-SGW is for Further Study.

5.2.5 Composite Gateway

Composite gateway: A logical entity composed of a single MGC and one or more MGs that may be reside on different machines. Together, they preserve the behaviour of a gateway as defined in H.323 and H.246 (this may include SIP servers and MSC servers in release 2000).

5.2.6 Media Gateway Control Function (MGCF)

This component in the R00 all-IP network is PSTN/PLMN termination point for a defined network. The functionality defined within MGCF should be consistent with existing/ongoing industry

protocols/interfaces that will satisfy the requirements.

- Controls the parts of the call state that pertain to connection control for media channels in a MGW.
- Communicates with CSCF.
- MGCF selects the CSCF depeneding on the routing number for incoming calls from legacy networks.
- Performs protocol conversion between the Legacy (e.g. ISUP, R1/R2 etc.) and the R00 all-IP network call control protocols (this is still under further study within the industry).
- Out of band information assumed to be received in MGCF and may be forwarded to CSCF/MGW.

Interfaces need to be further studied and defined.

5.2.7 Media Gateway Function (MGW)

This component in the R00 all-IP network is PSTN/PLMN transport termination point for a defined network. For the architecture option 2, the component is also used for interfacing UTRAN with the All IP core network over Iu.

The functionality defined within MGW should be consistent with existing/ongoing industry protocols/interfaces that will satisfy the requirements.

A MGW may terminate bearer channels from a switched circuit network (i.e., DSOs) and media streams from a packet network (e.g., RTP streams in an IP network). Over Iu MGW may support media conversion, bearer control and payload processing (e.g. codec, echo canceller, conference bridge) for support of different Iu options for CS services: AAL2/ATM based as well as RTP/UDP/IP based. *[Note: in the general R'00 architecture different core network transport technologies shall be possible for example: ATM, STM or IP.]*

- Interacts with MGCF, MSC server and GMSC server for resource control.
- Owns and handles resources such as echo cancellers etc.
- May need to have codecs.

In band signalling impacts to MGW and R00 all-IP network is for further study.

Functionality on the delivery of ring tone towards PSTN/ PLMN are for further study.

The MGW will be provisioned with the necessary resources for supporting UMTS/GSM transport media. Further tailoring (i.e packages) of the H.248 may be required to support additional codecs and framing protocols, etc.

For architecture option 2, the MGW bearer control and payload processing capabilities will also need to support mobile specific functions such as SRNS relocation/handover and anchoring (note that these functions are provided by the SGSN/GGSN in architecture option1 and are not required in the MGW). It is expected that current H.248 standard mechanisms can be applied to enable this. Solutions of how to use the H.248 generic bearer control mechanisms for mobile specific functions needs further studies.

Interfaces need to be further studied and defined.

5.2.8 Multimedia Resource Function (MRF)

- This component performs multiparty call and multi media conferencing functions. MRF would have the same functions of an MCU in an H.323 network.
- Responsible for bearer control (with 3G-GGSN and MGW) in case of multi party/multi media conference

♦ May communicate with CSCF for service validation for multiparty/multimedia sessions.

Handling of resources such as two stage dialling, announcements etc. are for further study.

Interfaces need to be further studied and defined.

5.2.9 MSC Server

MSC server mainly comprises the call control and mobility control parts of a GSM/UMTS R99 MSC.

The MSC Server is responsible for the control of mobile originated and mobile terminated 04.08CC CS Domain calls. It terminates the user-network signalling (04.08+ CC+MM) and translates it into the relevant network – network signalling. The MSC Server also contains a VLR to hold the mobile subscriber's service data and CAMEL related data.

MSC server controls the parts of the call state that pertain to connection control for media channels in a MGW.

5.2.10 Gateway MSC Server

The GMSC server mainly comprises the call control and mobility control parts of a GSM/UMTS R99 GMSC.

5.3 Description of Reference Points

5.3.1 Cx Reference Point (HSS – CSCF)

This reference point supports the transfer of data between the HSS and the CSCF.

When a UE has registered with a CSCF, the CSCF can update its location towards HSS. This will allow the HSS to determine which CSCF to direct incoming calls to. On this update towards the HSS, the HSS sends the subscriber data (application related) to CSCF.

For a MT call, CSCF asks the HSS for call routing information.

5.3.2 Gm Reference Point (CSCF – UE)

This interface is to allow UE to communicate with the CSCF e.g.

- register with a CSCF,
- Call origination and termination
- Supplementary services control.

5.3.43 Mc reference point (MGCF – MGW)

The Mc reference point describes the interfaces between the MGCF and MGW, between the MSC Server and MGW, and between the GMSC Server and MGW. It has the following properties:

- full compliance with the H.248 standard, baseline work of which is currently carried out in ITU-T Study Group 16, in conjunction with IETF MEGACO WG.

- flexible connection handling which allows support of different call models and different media processing purposes not restricted to H.323 usage.
- open architecture where extensions/Packages definition work on the interface may be carried out.
- dynamic sharing of MGW physical node resources. A physical MGW can be partitioned into logically separate virtual MGWs/domains consisting of a set of statically allocated Terminations.
- dynamic sharing of transmission resources between the domains as the MGW controls bearers and manage resources according to the H.248 protocols.

For architecture option 2, the functionality across the Mc reference point will need to support mobile specific functions such as SRNS relocation/handover and anchoring. It is expected that current H.248/IETF Megaco standard mechanisms can be applied to enable this. Solutions of how to use the H.248 generic bearer control mechanisms for mobile specific functions needs further studies.

5.3.4 Mh Reference Point (HSS – R-SGW)

This interface supports the exchange of mobility management and subscription data information between HSS and R99/legacy mobile networks. This is required to support All IP users who are roaming in a 2G network.

5.3.5 Mm reference Point (CSCF – Multimedia IP networks)

This is an IP interface between CSCF and IP networks. This interface is used, for example, to receive a call request from another VoIP call control server or terminal.

5.3.6 Mr Reference Point (CSCF - MRF)

Allows the CSCF to control the resources within the MRF

5.3.7 Ms reference Point (CSCF – R-SGW)

This interface allows CSCF to contact legacy network elements, e.g. 2G HLR, for location management (location update and subscriber data download), and call control (eg 2G HLR enquires for routing number (RN) for a roaming 2G user).

5.3.8 Mw Reference Point (CSCF – CSCF)

The interface allows one CSCF (e.g. home CSCF) to relay the call request to another CSCF (eg serving CSCF).

5.3.9 Nc Reference Points (MSC Server – GMSC Server)

Over the Nc reference point the Network-Network based call control is performed. Examples of this are ISUP or an evolvement of ISUP for bearer independent call control (BICC). In the all IP core network Nc reference point uses an IP based signalling transport. *[Note: in the general R'00 architecture different options for signalling transport on Nc shall be possible.]*

5.3.10 Nb Reference points (MGW-MGW)

Over the Nb reference point the bearer control and transport are performed. The transport may be RTP/UDP/IP or AAL2 for transport of user data. The bearer control over Nb is FFS, it may be based on RTP, H.245 or corresponding. *[Note: in the general R'00 architecture different options for user data transport and bearer control shall be possible on Nb, for example: AAL2/Q.AAL2, STM/none, RTP/H.245.]*

5.3.11 SGSN to Applications and Services

The interface from the SGSN to the SCP in the Applications and services domain is the interface defined for GPRS to support Charging Application Interworking.

5.4 Usage of MAP/CAP - Protocol stack below MAP / CAP – General considerations

Below MAP and CAP, the protocol stack within the All IP CN is as shown in Figure 5-5:

- SCCP and TCAP are used below CAP and MAP. Indeed CAP and MAP both rely on services provided by these underlying protocols (e.g., transaction capabilities, global title translation). Alternatives to providing the services of SCCP is for further study.
- The lower transmission layers are compliant with the IETF Sigtran protocol suite used to carry telecommunication signalling on top of an IP backbone.

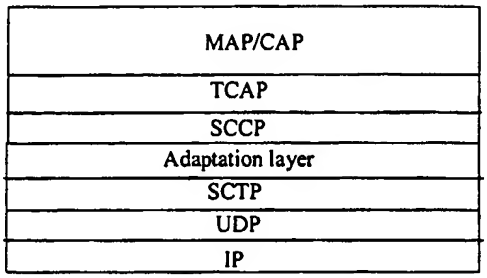


Figure 5-5: All IP R00 protocol stack for MAP/CAP

This protocol stack is used to carry MAP/CAP flows:

- inside the All IP CN, between nodes terminating the MAP/CAP: e.g. between HSS or SCP and the All IP CN functions (SGSN, MSC servers ...) handling Call / Session and needing to dialog with the HSS or SCP for user mobility management / subscriber data retrieval. This is for example the case of the Gr, Gc interfaces.
  It may also be envisaged (although this requires further study) to use MAP on Cx for user mobility management / subscriber data retrieval.
- in case of interworking with nodes not supporting this MAP/CAP over IP stack (e.g. 2G or UMTS R99 networks nodes) but needing a MAP / CAP dialog with a node supporting this MAP/CAP over IP stack. This protocol stack is used between the nodes terminating MAP/CAP and the R-SGW. The Mh interface shall use MAP over IP, for the roaming scenarios involving the R00 CS domain and GPRS (excluding the VoIP/Multimedia domain), hence this does not exclude other protocols being used. The use of MAP on the Ms is FFS.

6 QoS

The work currently being done within the S2 QoS Ad Hoc is reflected within TR 23.907 and the QoS section of TR 22.105. The R99 version of these specifications will support real time applications on a packet switched network which includes the ability of UMTS to transparently support multi-media applications that utilize the H.323 protocol. The all-IP architecture as described in this document defines the implementation of a call control function which can be based on either SIP or H.323 within the PLMN. Therefore, the R00 QoS work will include any changes required to

support QoS capabilities necessary for support of multi-media based on H.323 or SIP within the PLMN. Doing this is not expected to introduce any new QoS requirements at the UMTS bearer level.

In additon, please note that the desire also exists to have the all-IP architecture support multi-media applications that utilize the SIP protocol. However, QoS work related to the SIP protocol would only be undertaken within 3GPP when 3GPP itself undertakes the work to support the SIP protocol. However, we do anticipate that the QoS work centered on H.323 is directly applicable to SIP.

In addition, since the work on the all IP network includes EDGE support as identified by the ERAN architecture, the QoS work within the ERAN needs to be in alignment with QoS support within UMTS.

A preliminary review of the current version of TR 23.907 leads us to believe that it is largely sufficient to meet the objectives an all IP network. The review includes the following observations :

- TR 23.907 includes the specification of a QoS conversational class which includes voice. TR 23.907 identifies the fundamental characteristics of this class as:
- Preserve time relation (variation) between information entities of the stream.
- Conversational pattern (stringent and low delay)
- These characteristics apply whether voice is carried within a circuit or as packets. So there should be no need to modify the QoS classes currently defined.
- Implementing the H.323 call model within the PLMN is not expected to affect the R99 TR 23.907 identified QoS technical requirements, the overall architecture, nor the functions identified therein. However, a brief study will be necessary to verify this.
- The Radio Access Bearer Service attributes currently defined will need to be reviewed in light of an all IP network but minimal additions in this area would be expected for UMTS. However, for EDGE, work should be anticipated. Obviously the mapping from bearer to radio bearer is also affected.
- The issue of interworking the packet voice capable GPRS with other networks needs to be studied at least as it pertains to an acceptable voice delay budget.

7 Handover

Within this section, the topics to be studied and standardised to support handover for real time services in the PS domain have been identified. This section has investigated various handover scenarios, however the fact that the scenario has been studied here does NOT imply a requirement for the support of that scenario. The requirements for handover relating to the All IP architecture of R00 in UMTS will be determined by S1 as part of the R00 Service Requirement specification work.

7.1 SRNC Relocation within a UMTS R00 IP network

Within UMTS, work has already been undertaken to provide handover for real time PS domain services. The UTRAN does not distinguish between circuit and packet services, it simply provides for real and non-real time services, hence Intra RAN handover for real time services is available.

7.1.1 Support Required within ERAN

The goal of the All IP architecture is to provide a common core network for both UTRAN and ERAN. The specification of this work is outside the scope of this study, however, it is worth noting that the ERAN will need to support the following procedures:

- SRNC Relocation
- Mobile Assisted Network Controlled handover for the real time packet services.

7.1.2 All IP UTRAN to All IP ERAN Handover

The need to support this handover scenario is for FFS.

In this scenario, a CSCF will support terminals in both the ERAN and the UTRAN. The terminal will have access to the same Media Gateway from both RANs, hence the same media codec will be used in the network.

7.2 SRNC Relocation/Handover Between All IP and CS Domain/GSM

7.2.1 Requirement

The need to support these handover scenarios is for FFS.

The expected scenarios:

- Inter system handover, where target system does not support the necessary RT requirements for its packet domain (e.g. Inter system hand-over towards R97)

To fulfill this potential requirement, 2 solutions have been currently proposed (other may be studied). Any solution would face the following issues:

1. The MS has one or more PDP context for the signalling and the traffic. As the MS after the HO is handled by the CS domain, these PDP contexts need to be switched out? How to signal to the MS that its traffic is now handled by the CS domain? There is no H323 (H225 / H245) that could be used for such purpose?
2. Multimedia CC messages sizes may be larger than supported in the CS domain. The feasibility of transfering the Multi-media protocol messages on top of GSM CS signalling radio and A interface connections as well as on the MAPE interface needs to be investigated.
3. How does the SGSN determine that sessions involve the CSCF?

7.2.2 Solution with CSCF supporting MAP E

The following text considers the scenario when a UE has at least one session active which involves the CSCF.

On receipt of an SRNC relocation required message, the SGSN determines that the SRNC relocation results in a change of SGSN to one, which does not support the All IP services. One option is to force the serving RNS to hold the sessions until those involving the CSCF have been torn down, alternatively the SGSN needs to transfer the sessions not using the CSCF to the SGSN and the CSCF involved sessions to a 3G-MSC.

The signaling described below is based on the procedures for SRNC Relocation in 23.121. It shows that the use of the anchor MSC concept could be applied in order top maintain the Multimedia CC signaling to be tunneled back to the CSCF which acts at the "Anchor MSC". However, issues still arise as to how to remove the PDP context to terminate it within the network.

1. On receipt of "SRNC relocation Required Message" SGSN checks for
   - Do any sessions involve the CSCF? If so,
   - is the target SGSN an E-GSN?
2. SGSN signals to the CSCF that a handover to an MSC is required. I.e. sends "Forward SRNC Relocation" message
3. CSCF signals "Prepare SRNC Relocation" to Non-anchor MSC including the information received from the Source RNC.

4. Non-anchor MSC starts the Relocation process, treating the CSCF as the Anchor MSC. This allows the Multi-media client CC messages to be tunneled through the Non-anchor MSC to the CSCF.

5. The CSCF instructs the GW to prepare to transfer the traffic between the PSTN and the Non-Anchor MSC. I.e. to take the GGSN out of the path.

6. Successful switch of the bearers at the RNC, takes the SGSN and GGSN out of the path. The GTP tunnels are then released.

Open Issues:

1. How does the SGSN determine which sessions involve the CSCF?

2. How does the SGSN inform the CSCF to request a connection from the MSC via a MAP-E interface? In the case of the CSCF being in an external network, it may not be possible for the SGSN to know the CSCF address.

3. For a Mobile to PSTN call, the CSCF will need to signal to the GW to route PSTN traffic to the 3G-MSC

7.2.3 Inter-System handover using the ISHF

The mechanism described in this section, identifies a new functional element, the ISHF. This isolates MAP/E from the CSCF. Further work is required to identify if this approach, or the approach of supporting MAP/E on the CSCF (see section 7.2.2) should be adopted.

7.2.3.1 General

Based on the handover requirements given in Table 4-1, the following intersystem handover scenarios should be accommodated by the All IP architecture.

- UMTS R 00 IP network to/from 2G GSM network handover

These procedures listed shall not require change to the terminal.

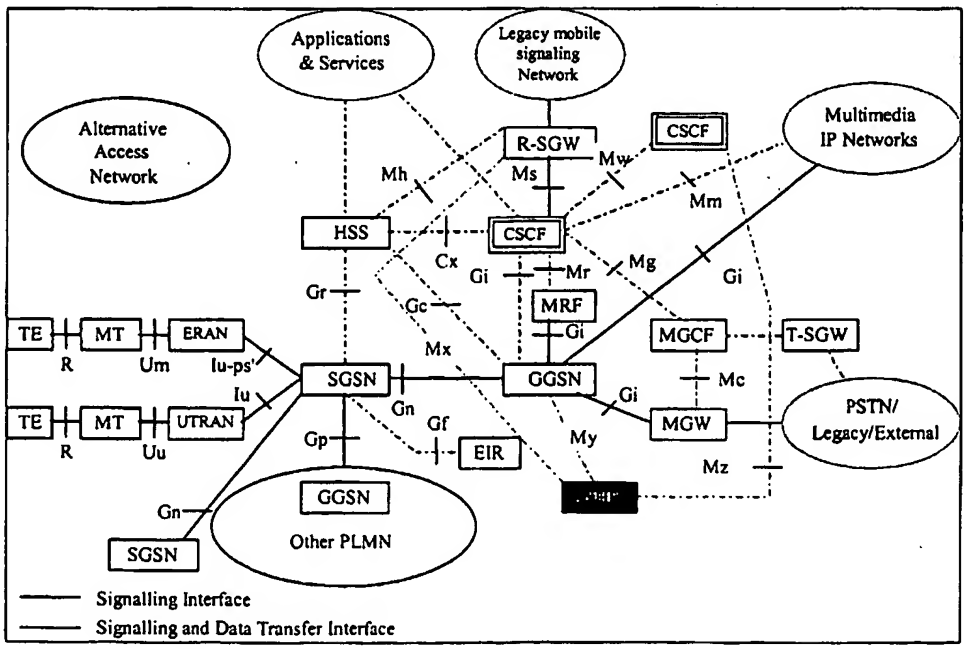


Figure 7-1: Support of InterSystem Handover

To support intersystem handover it is proposed that a new function is added to the architecture, called the InterSystem Handover Function (ISHF), see Figure 7-1.

The changes/additions to the baseline architecture given in Figure 7-1 include:

1. ISHF is the Inter-System Handover Function for handover between the UMTS R 00 IP networks and UMTS (PS, CS) networks and between UMTS R 00 IP networks and legacy networks. The ISHF is responsible for the handoff signaling procedures to another core network in addition to the establishment of a bearer connection between the source and target networks.
2. Mx is the interface between ISHF and the R-SGW. This interface is MAP/E and is used to signal handoff messaging between networks.
3. My is the interface between the ISHF and the GGSN. This interface relays handover related Iu signaling between the UTRAN and the ISHF. Note this is interface is tunneled through the SGSN.
4. Mz is the interface between the ISHF and the CSCF. This interface is used setup bearer resources between the source and target networks for inter-system handover.

It is assumed that the R-SGW function will interwork the UMTS R 00 IP handover procedures to the handover procedures of the source or destination network, that is the ISHF resides within the R-SGW. It may be desirable to create a separate function that performs interworking between core network protocols, this is for FFS.

7.2.3.2 UMTS R 00 IP network to/from 2G network handover

This example shows how handover (Hard Handover) is performed from UMTS R 00 IP network to a legacy GSM network. This demonstrates the signaling required between the networks and assumes a trunk circuit bearer between the networks. Other bearer connection schemes are possible, but not addressed in this example. (Note that all the air interface messages are not shown for clarity in the diagram. In addition, the MGCF and T-SGW are shown as a combined node and the messaging between these functions are omitted for clarity.)

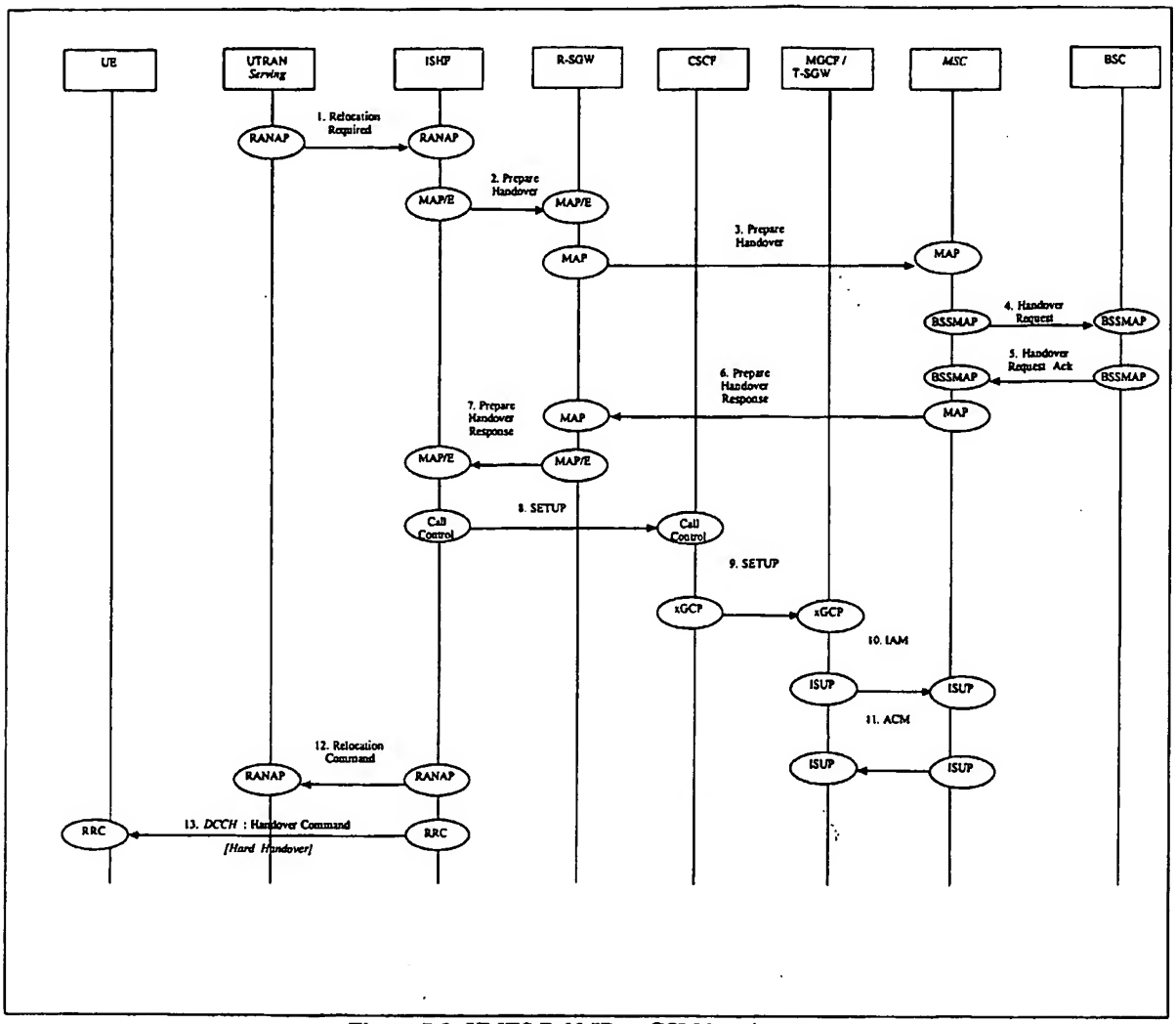


Figure 7-2: UMTS R 00 IP to GSM handover

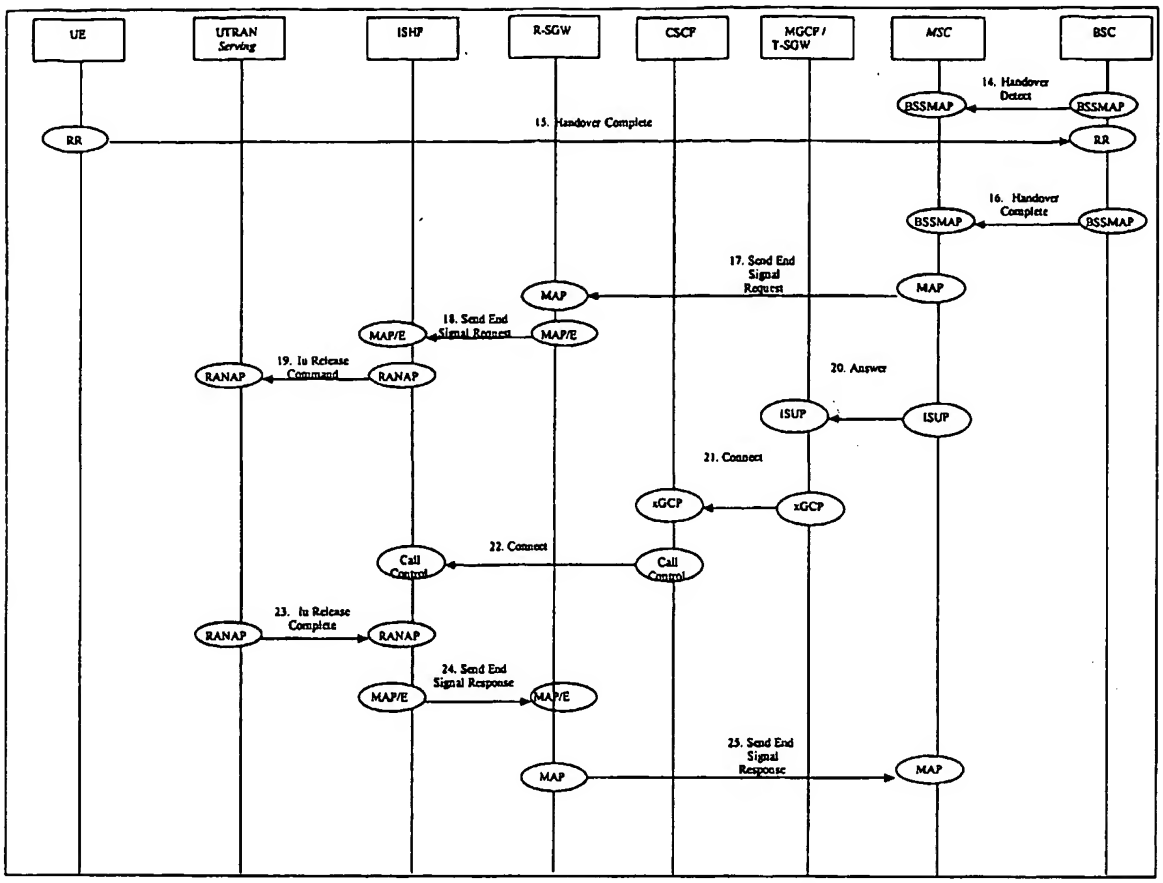


Figure 7-3: UMTS R 00 IP to GSM handover (continued)

UMTS R 00 IP ⇒ GSM handover

1. Upon detection of a trigger SRNC sends RANAP message **Relocation Required** to the ISHF.
2. The ISHF will send the MAP/E **Prepare Handover** to the R-SGW.
3. The R-SGW Interworks (if required) the **Prepare Handover** to the appropriate network protocol (in this case GSM MAP) and sends the message to the other network (MSC)

Note: Steps 4&5 follow the normal GSM procedures and are shown only for clarity.

6. Once initial procedures are complete in GSM MSC/BSS the MSC returns MAP message **Prepare Handover Response** to the R-SGW.
7. The R-SGW converts (if required) to the MAP/E protocol and sends the resulting **Prepare Handover Response** message to the ISHF.
8. The ISHF initiates procedures to establish bearer resources between the networks. In this case a trunk circuit is established. The ISH sends a **CONNECT** message to the CSCF to initiate a call to set up the bearer.
9. The CSCF sends a **CONNECT** to the circuit gateway (MGCF, T-SGW shown combined for simplicity) to establish an outgoing call to the MSC.
10. The circuit gateway sends the ISUP **IAM** message to the MSC.
11. The MSC responds with the ACM message.
12. ISHF responds to the initial request from SRNC by sending RANAP message **Relocation Command** to the SRNC.
13. Via existing RRC connection, SRNC sends RRC message **Handover Command** (Hard Handover) to the UE.

Parameters: Handover type.

Note: Procedures related to synchronization etc. to GSM BSS are not shown.

Note: Step 14-16 follow normal GSM procedures and are shown only for clarity.

17. Detection of the UE within the GSM coverage results in the MSC sending MAP message **Send End Signal Request** to the UMTS R 00 IP network (R-SGW)

18. The R-SGW forwards the **Send End Signal Request** to the ISHF.

19. ISHF initiates release of resources allocated by the former SRNC (**Iu Release Command**).

20. **ISUPAnswer** is sent from the MSC to the Circuit Gateway.

21. **Connect** is returned to the CSCF function

22. **Connect** is relayed back to the ISHF.

23. Previously allocated bearer resources are released within UMTS (e.g. using RANAP and ALCAP protocols [ALCAP not shown]) (**Iu Release Complete**).

24. Procedure is concluded from UMTS point of view by ISHF sending MAP/E message **Send End Signal Response** (this message is not sent until the end of the call).

25. The R-SGW will send the MAP **Send End Signal Response** to the MSC.

7.3 Areas for Further Study

The following areas may require further study.

- Bearer set-up/control between networks during handover
- Anchoring bearer in the UMTS R 00 IP network
- MAHO support
- Inter-RNC Soft handover
- Inter RAN to RAN of same type streamlining
- Inter RAN to RAN of different type streamlining

8 Radio Aspects

Note: This section requires support from the RAN group.

8.1 General

1) CN – RAN interface definition
   (a) Functional split between CN and RAN- new radio access network called EGPRS Radio Access Network (ERAN) is considered. The interface between the radio access network such as ERAN or UTRAN and the CN needs to be defined/extended and should allow different air interface technologies to access to the CN. The detailed functional split between CN and RAN needs further investigation.

(b) Impact on existing GPRS/EGPRS/UTRAN implementations and deployments

(c) Migration scenarios-2G to release 99 (BSS is not IP-based) to release 2000 (all IP-based network)

(d) Protocol stack evaluation (including evaluation of control and user planes from CN to both RAN and MS)

2) ERAN architecture (Refer to SMG2)

(a) ERAN reference model-network entities, protocol stacks and logical or functional elements

(b) Functional split between elements

(c) Definition of interaction between elements

(d) Impact on existing GPRS/EGPRS deployments (BSC, PCU, and BTS) and mitigation strategies

3) UTRAN architecture extensions

(a) Identification of required extensions

4) Realtime Handover for Packet Domain

(a) ERAN issues (Refer to SMG2)

(b) UTRAN issues

5) QoS support

(a) Evaluation of S2 QoS Ad Hoc progress for real-time data support

(b) Signalling mechanism

(c) CN issues

(d) Alignment of GPRS with UMTS QoS

(e) Realization of QoS on radio link

6) (E)GPRS Radio issues (Refer to SMG2)

(a) Real-time support including handover and QoS

(b) Spectrum efficiency/performance (e.g. statistical multiplexing and source/channel coding)

(c) RLC/MAC enhancements

(d) The effect of various deployment scenarios (e.g. spectrum availability) and traffic mix, such as voice and data, on spectrum efficiency should be considered.

7) UTRA Radio issues

(a) Real-time packet data support including handover and QoS – validation and possible enhancement

(b) Radio efficiency/performance (e.g. statistical multiplexing and source/channel coding)

(c) RLC/MAC enhancements if needed

(d) The effect of various deployment scenarios (e.g. spectrum availability) and traffic mix, such as voice and data, on spectrum efficiency should be considered.

Note: The RLC/MAC line items about (sections 6(c) and 7(c)) may include

- Enhanced for radio resource allocation.
- Radio access bearer definitions (i.e., define for the various traffic classes the path through the protocol stack and the bearer to be used).
- Flow classification (e.g. mapping of user traffic onto appropriate radio access bearer)
- Enhancement of EGPRS protocol to support real time service and QoS management (e.g. Fast channel allocation schemes).

8.2 Airlink Optimisation for Real-Time IP

8.2.1 Introduction

In the all-IP architecture, a fundamental objective is to support IP-based real-time and non real-time traffic for a mobile terminal while achieving spectral efficiency and error robustness. In the case of real-time voice, spectral efficiency and error robustness have a performance baseline coming from the current cellular systems. There is also a baseline in voice quality. It is natural to expect that the all-IP architecture has to meet this existing baseline for voice services. The question is then how to meet the objectives of spectral efficiency and error robustness and the existing baseline for real-time voice when the all-IP paradigm is applied to cellular systems.

For IP-based real-time multimedia, RTP protocol is predominantly used on top of UDP/ IP. The size of the combined IP/UDP/RTP headers is at least 40 bytes for IPv4 and at least 60 bytes for IPv6, while the voice payload is short, typically shorter than the IP/UDP/RTP header. Clearly, if the headers were sent "as is" over the air interface to conform to the pure IP paradigm, it is not possible to meet or even get close to the baseline spectral efficiency of existing circuit voice. Some header adaptation technique is required, whereby a transformation is applied to the IP/UDP/RTP headers to reduce their size before transmission on the air interface, and the reverse transformation applied after crossing the air interface, to restore the original header size and values. Reduction of the header size is done by removing redundancy in the originally coded header information and/or removing header field information and thereby losing functionality. Impact on transparency and robustness to errors have to be fully understood in order to design the appropriate adaptation techniques (Transparency for a given header field is defined as the property whereby the value after transformation/reverse transformation is the same as in the original header).

This section explores the range of possible adaptation techniques and proposes two adaptation techniques, header stripping and header compression. The two techniques must be further studied as regards error robustness, voice quality and IP transparency.

8.2.2 user plane adaptation

In the following we refer to the functionality that does transformation/reverse transformation as User Plane Adaptation (UPA), and explore the range of possible adaptation techniques, along with their pros and cons.

8.2.2.1 Full opacity (no adaptation)

The UPA has no knowledge of the internal structure of the headers or payload, and no transformation is done on the IP/UDP/RTP headers which are sent in full over the air interface. Error protection is applied evenly to all the bits in the header, and evenly to all the bits in the payload. The header part will likely require stronger error protection than the payload, since a header loss will require to discard the corresponding packet, and no error concealment or mitigation can be applied to the header. This technique achieves full transparency, which allows to support protocols such as IPSEC on an end-to-end basis. An obvious con is the high overhead caused by the headers, which results in very poor spectrum efficiency.

8.2.2.2 Payload opacity (header adaptation only)

In this case, the UPA only needs to know the internal structure of the IP/UDP/RTP header but not of the payload. Only the headers are adapted, either by header compression or header stripping.

8.2.2.2.1 Header compression/decompression

IP/UDP/RTP headers are compressed before transmission over air interface and decompressed at the receiving end. Like before, headers require stronger error protection than payload. The most wellknown header compression algorithm is the Van Jacobson algorithm (RFC 1144, Compressing TCP/IP Headers for low speed serial links). In general compressed headers are more vulnerable to errors than uncompressed headers. The current standardised algorithms has therefore proven to be less efficient over lossy links such as a radio interface.

The benefit of compressing the IP/UDP/RTP header is nevertheless obvious as it significantly reduces the required overhead per packet.For an efficient header compression scheme, the IP/UDP/RTP headers can be compressed down to 2 bytes.

New header compression schemes, adapted to cellular radio link reliability characteristics, will be developed in the future. Such schemes adapted to radio environment may be able to compress the IP/UDP/RTP headers down to 2 bytes. One example among others would be a scheme currently being proposed for development in the IETF in which the compressed header carry a checksum computed over the header before compression. This provides a reliable way to detect and repair errors and increases error robustness.

A general drawback with most header compression schemes are incompatibility with end-to-end security (IPSEC) and bandwidth management, since compressed headers have variable size.

Requirements and evaluation criteria for header compression schemes to be used over the radio interface is summarised in the table 8-1.

Figure 8-1 shows a conceptual diagram of header compression used in conjunction with the lower layers in cellular. Voice is used as an example. The lower layers may perform interleaving and channel coding. For simplicity, the effect of interleaving and channel coding on the bit stream transmitted over the air interface is not shown. The effect of possible link level multiplexing with other traffic streams is not shown either. There is an MS-based UPA point and a network-based UPA point. The MS-based UPA acts as header compressor and header decompressor for the uplink and downlink respectively, while the network-based UPA acts as header decompressor and header compressor for the uplink and downlink respectively.

The requirements for header compression are described in the table below. In each case, some justification for the requirement is also provided.

#	Focus Area	Requirement	Justification
1	**Performance / Spectral Efficiency**	Must provide low relative overhead (as defined in [1]) under expected operating conditions	In general, a primary goal is high spectral efficiency. Reduction of overhead has direct impact.
2	**IPv6 or IPv4**	Must include support for IPv6 and IPv4	Ipv4 and Ipv6 terminals are expected to coexist for some time
3	**Ubiquity**	Must NOT require modifications to existing IP (v4 or v6), UDP, or RTP protocol stack implementations	Enables use of current devices/services which employ generic IP/UDP/RTP stacks.
4	**Cellular HO**	Must support the cellular handoff operation, in an efficient manner; All fields must be transparent to the HO process, i.e, are exactly regenerated subsequent to handoff.	Target application is for adaptation of the user plane on cellular air interfaces; therefore this operation must be supported. Efficiency requirement is due to potential impacts on spectral efficiency and voice quality if HO is not properly handled.
5	**Integrity**	***The header compression process must be lossless***	Would like to maintain the end-to-end integrity of IP
6	**Error Propagation**	Error propoagation due to header compresion should be kept to a absolute minimum or avoided if at all possible.; *error propagation* is defined as the loss of packets subsequent to the one where the error actually occurred, even when those subsequent packets contain no errors	Error propagation results in lower spectral efficiency and lower voice quality; this is a serious problem for existing schemes such as [5].
7	**Delay**	Must operate under all expected delay conditions; header compression process must not contribute significantly to system delay budget	The user may be in different types of environments with different characteristics; additional delays will have adverse effects on conversational voice
8	**Packet Loss**	Must operate under all expected packet loss conditions; prefer that header compression efficiency is as independent of packet loss rate as possible	The user may be in different types of environments with different characteristics

9	**Media Supported**	Must function regardless of media type in RTP payload (in general, there is NO *required* knowledge of payload)	The algorithm should be applicable to any type of RTP/UDP/IP data flow; note that this does not preclude optional optimizations for certain media types
10	**Independence with respect to call type**	Must function for mobile-mobile and mobile-landline calls; performance in each case should be comparable to existing cellular (in terms of both quality and spectral efficiency)	Both types of calls will occur in All-IP cellular systems; each is equally important

Table 8-1: Requirements for Header Compression

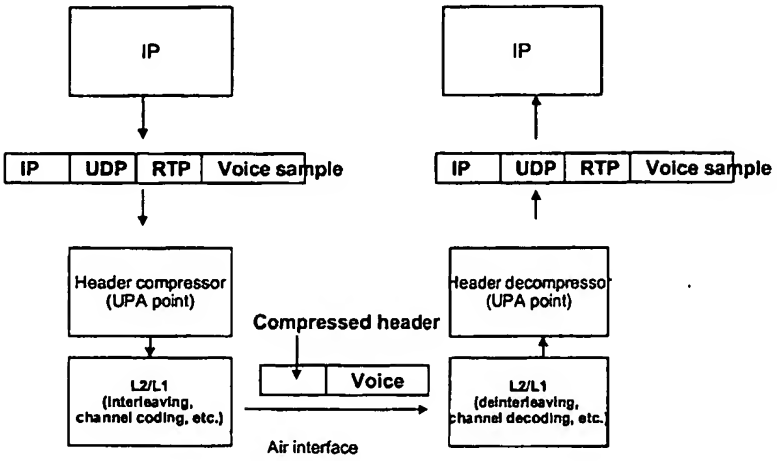


Figure 8-1: Header compression

8.2.2.2.2 Header stripping/regeneration

IP/UDP/RTP headers are stripped before transmission over air interface and regenerated at the receiving end. Essentially only the payload is transmitted, but some additional header-related information needs to be transmitted to enable the header regeneration. The degree of header transparency achieved is variable, depending on the amount of header-related information that one wants to transmit. No header error protection is needed when the header information is completely removed. However, the necessary information for header regeneration requires a header, at least for some packets. When the payload has constant size, bandwidth management issue is virtually eliminated since the payloads can be carried on a constant bit rate channel. The constant bit rate channel also eliminates QoS (delay and jitter) problems. As before, end-to-end security cannot be accomodated.

Figure 8-2 shows a conceptual diagram of header stripping used in conjunction with the lower layers in cellular. Voice is used as an example. The lower layers may perform interleaving and channel coding. For simplicity, the effect of interleaving and channel coding on the bit stream transmitted over the air interface is not shown. The effect of possible link level multiplexing with other traffic streams is not shown either. There is an MS-based UPA point and a network-based UPA point. The MS-based UPA acts as header stripper and header regenerator for the uplink and downlink respectively, while the network-based UPA acts as header regenerator and header stripper for the uplink and downlink respectively.

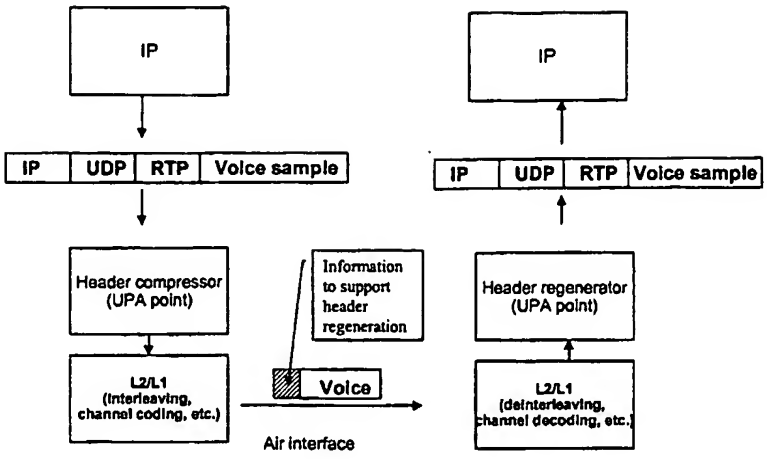


Figure 8-2: Header stripping

8.2.2.3 No opacity (full adaptation)

The UPA knows the structure of the headers and the payload. Headers can be compressed or stripped. In addition, payload transmission is optimised by techniques such as unequal bit protection, channel and error coding optimised for the payload structure, etc.

8.2.3 Application to all-IP network

The all-IP network is expected to provide real-time bearer services intended to carry

- Basic conversational voice (service equivalent to voice in current cellular)
- Real-time Multimedia (includes voice which is seen as a component of multimedia)

8.2.3.1 Basic voice

For basic voice, the emphasis is on meeting and if possible exceeding the baseline of traditional cellular in terms of spectrum efficiency, error robustness and voice quality. Traditional cellular systems achieve that baseline by using well known techniques such as unequal bit protection, channel and error coding optimised for the payload, etc. In addition, speech frames do not incur any header (in the IP sense) overhead. In the all-IP system, we propose to define a "basic voice" bearer tailored for conversational voice and possible other media with the same characteristics.

The basic voice will use payload optimisation and unequal bit protection of the payload similar to traditional cellular. Packing more speech frames into one packet will improve the relative overhead, but at the expense of added delay, which negatively impacts voice quality. Transmission of header-related information and/or compressed header will require strong error protection.

Two options eexist to achieve required characteristics for Basic Voice:

Header Stripping: At a minimum, header stripping for basic voice will have to achieve transparency for the static IP/UDP/RTP fields (those that do not change during the call) and the RTP time stamp and RTP sequence number. This bearer corresponds to the full adaptation case above with header stripping.

Header compression adapted to radio characteristics: By using a robust header compression scheme the overhead per packet is reduced to 2 bytes. Also this case shall correspond to the full adaptation case above with header compression.

Additional optimisation techniques may be contemplated to further improve the spectrum efficiency.

The two options and specific algorithms shall be evaluated according to the criteria of table 8-1 of chapter 8.1.2.1.

8.2.3.2 Real-Time multimedia

Real-time multimedia is a new service that does not exist in traditional 2G cellular systems. A new bearer is proposed. For that bearer, transparency for all the IP/UDP/RTP fields is crucial. Under the transparency constraint, we want to optimise spectrum efficiency and error robustness, but unlike voice, there is no baseline to be used as target. The transparency objective naturally leads to choosing header compression as the user plane adaptation. Payload will have some error protection and compressed header will have even stronger protection. The ability to provide unequal bit protection of the payload also for this service needs to be studied. This bearer corresponds to the header adaptation only case above with header compression. Specific algorithms applied to Real-Time multimedia shall be evaluated according to the criteria of table 8-1 of chapter 8.1.2.1.

8.2.3.3 Pure IP

The Pure IP service can be provided to accommodate end-to-end protocols such as IPSEC. In order to achieve this accomodation, the bearer does not do any adaptation and corresponds to the "No adaptation" case above. Header adaptation may also apply for Pure IP. Specific algorithms for header adaptation shall be evaluated according to the criteria of table 8-1 of chapter 8.1.2.1.

8.2.4 Conclusions

IP/UDP/RTP packets require adaptation to the radio link to meet the spectrum efficiency and error robustness requirements of cellular systems. It shall be investigated if a single scheme can simultaneously and fully meet the above requirements and IP transparency. An alternative to a single scheme is a gradation of schemes tailored to the particular

type of application. Applications may then use different kinds of bearers optimized for their particular current needs.

real-time bearers can be categorised in Seen from the header compression point of view basic voice (BV) and real-time multimedia (RTMM). BV bearer is intended to carry voice, as a service equivalent to the one in traditional cellular systems. RTMM bearer is intended to carry generic multimedia traffic, which can include voice. In addition, a pure IP service may be contemplated for support of applications, which require full transparency.

- BV will use header stripping or header compression with unequal bit protection in the payload.
- RTMM will use header compression with equal bit protection in the payload. The possibility to support unequal bit protection in the payload shall be investigated.
- Pure IP service will support data transport without transforming the header. Header compression shall also be possible for Pure IP. Pure IP uses equal bit protection in the payload.

In all cases, header (if present) requires strong error protection.

9 Call Control

9.1 Terminology for Call Control

The terminology in this section is that terminology used that is new or has been changed from that defined for R99. The terminology defined in this section has not been the object of a real debate and hence cannot be considered as agreed. This section needs to be aligned with the terminology used in R99. R99 terminology

should be used unless a new or changed concept is introduced. Changes to terminology defined by S1 require S1 agreement.

This section defines a common set of terms on which the present document is based on. The following list of terms is the first attempt to define some terminology.

The terminology defined here have not been matched with the existing 3GPP terminology and this matching will need to be done. Moreover, 3GPP has not defined all the terms that are needed for an all IP based network yet.

1 **Access Profile:** contains subscription profile information relevant to a specific bearer network. As an example, E-GPRS profile plus the radio bearer features (e.g. QoS) to which the user can access is an Access Profile. Access specific roaming information for access to and from legacy networks is a part of the access profile . As an example for CS terminals, Access Profile contains information on the allowed LA, on security data for authentication

2 **Release 2000 all IP networks Service Profile:** contains service subscription data relevant to the Release 2000 all IP network services the user has subscribed to. As an example, Service Profile contains user's identifier, user's aliases, user's temporary location information (e.g. pointer to the current Serving Domain), indication of the multimedia services and capabilities the user has subscribed to, service triggers, status of supplementary services, etc.

3 **User Profile:** is a combination of one or more Access Profiles and zero, one or more Release 2000 all IP networks and Roaming Service Profiles. The User Profile is maintained in the HSS. (FFS)

4 **PDP Flow:** it is a PDP context without the restriction that a different IP address has to be assigned to different PDP contexts. Differentiation between PDP flows is based on protocol type (e.g. TCP, UDP etc.) and port number. (FFS)

5 **Bearer Network:** it is a set of network elements that provides a user with means to connect to a serving/home domain to use services or facilities of the network the user is roaming in and gain access to the home domain or other service networks. Examples of bearer networks are:

- E-GPRS plus one or more different RANs;
- Cable Access Network, etc.

6. **Domain:** A domain is a logical association of network elements. A domain may contain any number of HSS, CSCFs and MRF. A domain can be Home Domain for some users (those whose subscription profile is stored in the HSS in that specific domain) and Serving Domain for other users (those whose subscription profile is stored in the HSS in a distinct domain). A domain can connect to a multitude of bearer networks. (FFS) The purpose of introducing the domain concept in release 00 is to enforce access Independence in the core, support of NAI addresses, scalability, additional services and servers expected (ex. Email/CSCF interworking), allow the use of DNS and directories for translations. Domain's already provide the glue for many useful services and functionalities. A domain is a logical realm that indicates a system or zone. A domain is used to associate services and servers with a common identifier for translation purposes. A domain is used as a key into databases in order to obtain the network addresses of nodes for respective services associated with the domain. The actual location of nodes supporting these services is not restricted in any way by a domain. The term domain used here refers to the DNS definition of domain.

Open Issue: The use of the term, domain, home domain and serving domain requires further clarification and analysis.

7 **Home Domain:** it is a domain that contains an HSS. In particular, the Home Domain of a user is the domain containing the HSS that stores the user profile. Home Domain may or may not contain the Home CSCF, the MRF or service logic. Home Domain:

- provides and maintains the user and user's subscription data in a HSS for the user belonging to the domain;
- supports also access independent users and users profile such as browser bookmarks and phone lists;
- user provides and updates the currently visited serving domain with user's profile;
- store routing and mobility information that enables service delivery to users and users roaming outside and inside the home network;
- maintains roaming agreements and service level agreements with other networks;

- Home Domain is seen as the initial termination point from the originating network when it contains the Home CSCF;
- may collect and consolidate charging data.

Other functions are FFS.

8 **Serving Domain**:

- stores roaming profile as received from the home domain
- provides services or access to services in the home domain (as per terminal capabilities and service level agreement if different operator domain) with the same "look, touch and feel" as much as possible;
- stores routing and mobility information that enables service delivery to users roaming in the service domain;
- optionally collects data for billing and statistics;
- can provide local services such as location based services and information (e.g. advertisements, operator announcements to local events, etc.);
- provides optional resources such as conference devices, multimedia call control etc. (Resource could be provided in home network)
- Home Domain can act as Serving Domain when the user is registered in the home domain. (FFS)

9 **Release 2000 all IP networks**: A Release 2000 all IP networks comprises of the following logical components:

- One or more domain(s);
- Any number of bearer network(s);
- Connectivity to one or more MGCFs and MGWs.
- Zero, one or more MSC/GMSC Servers

10 **Home Network (HN):** considering a specific user, the Home Network is made of zero, one or more serving domains, any number of bearer networks, zero, one or more MGCF/MGW, and the home domain for that specific user. (FFS)

11 **Serving Network (SN):** considering a specific user, the Serving Network is made of zero, one or more serving domains and one or more bearer networks, zero, one or more MGCF/MGW, and does not contain the home domain of that specific user. (FFS)

12 **Service Logic Domain**: includes the following functions:

- contains existing telecom service capabilities (i.e. SCP for IN);
- contains WAP type capabilities for Web-based services;
- allows easy access to services by the users (notifications of new services, activation/deactivation of services in the network, capability for payment and upgrade for new services, etc.);
- updates profile data in the home domain in the event of modifications by the user or by the provider of the Service Logic Domain operator;
- provides access to or capabilities to reach other Service Logic Domain;
- some location based services can also reside in the Service Logic Domain.

The Service Logic Domain could have a tight coupling with the Home Network in order to manage/charge/provide application service capability set uniformly to the users. In particular, if CSCF handles service triggers the Service Logic Domain and Home Domain may be providing both the user's user profile and subscription profile to CSCF in a

co-ordinated/transparent fashion. The Service Logic Domain could also be stand-alone, i.e. independent from the network location. (FFS) It is for further study whether or not the CSCF can be logically associated with the service logic domain.

13 **Home User**: is a user of the home network having a subscription in the home domain. A user is considered a home user when the user is located in a serving domain in its home network. The user may or may not be located in their home domain.

14 **Roaming User**: is a user roaming outside its Home Network and being served in a Serving Network. (FFS)

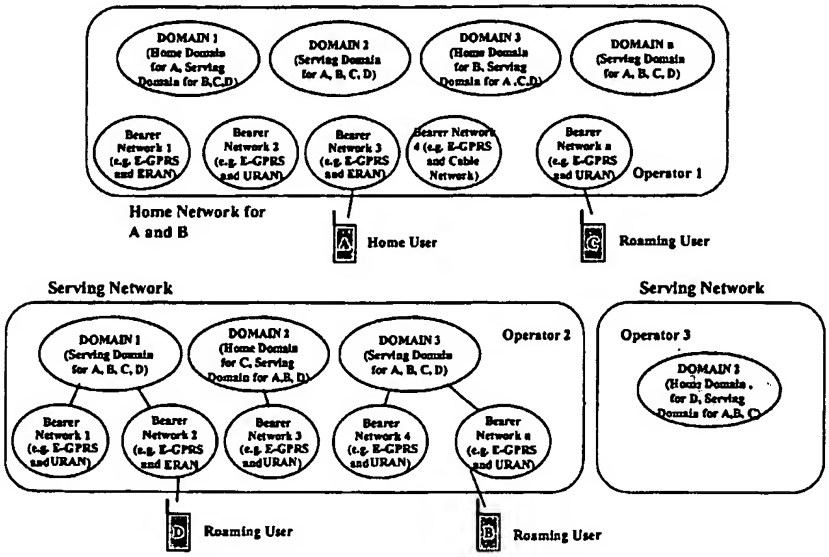


Figure 9-1: Modelling of the network in domains

15 **Legacy Network**: a legacy network can be:

- SS7 based networks (e.g. PLMN, PSTN and ISDN) as well as CAS based networks;
- GPRS networks. (FFS)

16 **Multimedia IP Network**: it is an external IP network with support of real-time multimedia services (using H.323 and/or SIP protocols), and includes SIP/H.323 network elements and terminals, and possibly gateways to interface with Release 2000 all IP networks. (FFS)

17 **Serving CSCF (S-CSCF)**: it is a CSCF in the Serving Domain with which the user is registered and that is providing the services depending on the Service Profile(s) obtained from the Home Domain of the user. (FFS)

18 **Home CSCF (H-CSCF)**: the Home CSCF is a CSCF in the Home Domain or Private Domain. The Home CSCF is associated to a user at the subscription time and, if the user is identified by aliases that might require translation to an IP address (e.g. logical names translated by DNS), the transport address of the Home CSCF will be provided as translation of the aliases. (FFS)

9.2 Assumptions

The following assumptions have been considered in the development of the roaming models described in the present version of the document.

1 The addressing requirements and mechanisms will be based on the requirements and mechanisms identified by 3GPP in 3G TR 22.975 and 3G TS 33.003.

2 Call admission/denying and call re-routing will be considered. Details of call admission (e.g. authentication and QoS) will not be discussed here.

3 Re-routing of incoming voice or data communication requests that are addressed to the user's directory number during periods of realignment of the national numbering plans will be considered. Probably, a non final solution (e.g. re-direction of calls based on databases) will be provided.

4 A specific PDP flow (called Signalling PDP flow), distinct from PDP flows carrying media PDUs, is adopted to carry signalling between UE and CSCF (e.g. call set-up, in-call signalling such as flash requests). The signalling PDP flow does not need to have an IP address different from the one of the PDP flows carrying the media.

5 The user profiles are stored in a permanent way into a database/server in the Home Domain.

6 In case of roaming the user profile in the home network must be interrogated by the serving network at least at registration and part of user profile could be temporarily stored into the serving network.

7 The roaming architecture will be optimised with the assumption that IP addresses will be allocated dynamically.

8 No new requirements would be placed on the legacy (e.g. PSTN, 2G PLMN) and Multimedia IP Networks to interconnect with the Release 2000 all IP networks. Release 2000 all IP networks would have to ensure inter-operability with the existing legacy and Multimedia IP Networks. In case a 2G HLR (e.g. GPRS HLR) is re-used to hold data for Release 2000 all IP network users, the 2G HLR will be upgraded to support Release 2000 all IP network user and its interfaces might need to be upgraded.

10. An MS registration procedure consists logically of a bearer level registration (e.g. GPRS attach) and, if so specified/allowed by the MS subscription profile, an application level registration. MS registration is performed, as an example, at MS power up.

11. The bearer level registration procedure entails registration with the GPRS nodes, according to GPRS-derived procedures.

12. The application level registration procedure entails registration with a serving CSCF/MSC server in order to inform the CSCF/MSC server of the MS presence and to allow the CSCF/MSC server to retrieve the user information from the HSS.

13. Bearer level registration and application level registration are considered to be two separate procedures. Bearer level registration completeion may trigger in the UE a signalling PDP flow activation as a possible mean of supporting the application level registration procedure.

14. The MS location is tracked at the GPRS level using GPRS mobility management, and user mobility is tracked at the application level through a specific procedure aimed at updating the user information maintained by the CSCF / MSC Server and, possibly, location information held by the HSS."

15. HSS keeps track of the user mobility in terms of the current CSCF/MSC server or Serving Domain.

16. Support of multiparty voice and data communication sessions (including the capability for the user or service logic to dynamically add or delete users from an active communication session) is not considered in the present version of this document.

17. Roaming agreements (static or dynamic) between the co-operating operators must exist.

18. Service Level Agreements (SLA) between network operators are defined (statically or dynamically) to ensure consistent level of services (e.g. end-to-end QoS, security etc.).

19. All network components that require address analysis and address translation for routing of terminating calls (e.g. CSCF, MSC server, MGCF) are capable of doing so or has access to consistent translation databases or to the HSS in order to resolve routing/call termination issues.

20. O&M functions exist to connect various components of a Release 2000 all IP network to each other, making it possible for each component to know how to address/access any other component within the network domain.

21. O&M functions exist to make network provisioning profiles (e.g. 800 triggers, tone information) available in each domain whenever.

22. In order to obtain user-plane optimisation, the serving GGSN will be preferably located in the serving network.

23. Service profile (subscription and activation status) and service triggers are either maintained by the HSS and/or updated by the HSS towards the CSCF/MSC Server.

24. CSCF-routed call signalling is assumed for real-time services.

25. Every CSCF is associated to one or more MRF, and every MRF can be controlled by more than one CSCF. For sake of clarity, if H.323 terminology is used to describe the MRF, MCU is part of the MRF. MRF could be located in the home network (when the Home CSCF is controlling the current call) or in the serving network (when the serving CSCF is controlling the current call).

26. Some of the functions of the A Home CSCF is are needed in all the roaming scenarios in order to support:

    - incoming calls addressed to a DN from other Release 2000 all IP networks or Multimedia IP Networks with optimised routing (i.e. calls not routed through PSTN);
    - incoming calls from other Release 2000 all IP networks or Multimedia IP Networks originated with a LN (Logical Name);
    - implementation of supplementary services and Incoming Call Screening-like functions for terminated calls (e.g. Call Forwarding Unconditional).

9.3 Roaming Within All IP networks

In the follow, a set of roaming scenarios is described.

***Editor's Note:** please note that the network interfaces and the names shown in the diagrams from 6.4 to 6.7 may not always be correct.*

9.3.1 Call Model

The call model described by the following statements has been adopted in the present document:

- Calls from/through PSTN are routed to an MGCF with connectivity to the Home Network corresponding to the dialled DN.
- Calls from a Release 2000 all IP network to a different Release 2000 all IP networks originated with a DN can be optimised (i.e. not routed to PSTN) only if the originating Release 2000 all IP networks has knowledge of the numbering plan of the destination Release 2000 all IP networks and can route the call directly to the destination Release 2000 all IP networks without leaving the IP domain. Further optimisations could be possible if the network where the call is originated has also access to the HSS of the network to which the call is destined and corresponding to the dialled DN. The same applies to calls from a Multimedia IP Network to a Release 2000 all IP network originated with a DN. (FFS)
- Assuming scenario 1, for incoming calls (i.e. mobile terminated calls), the call setup request always arrives at the ICGW which interrogates the HSS, implements Incoming Call Screening and relays (without performing any call control function) the request to Serving CSCF / MSC Server. If the originating Release 2000 all IP networks had the capability to interrogate HSS of the destination network, it would be possible to address the call setup request to the Serving CSCF directly.
- Regarding originating calls, the call request is handled by the Serving CSCF/MSc server (when present), or the Home CSCF when a Serving CSCF is not present.

This section covers only PS services.

9.3.2 Scenario 1, Traditional Model

The following pictures show respectively the roaming scenario 1 applied to roaming inside a single network and applied to roaming between networks.

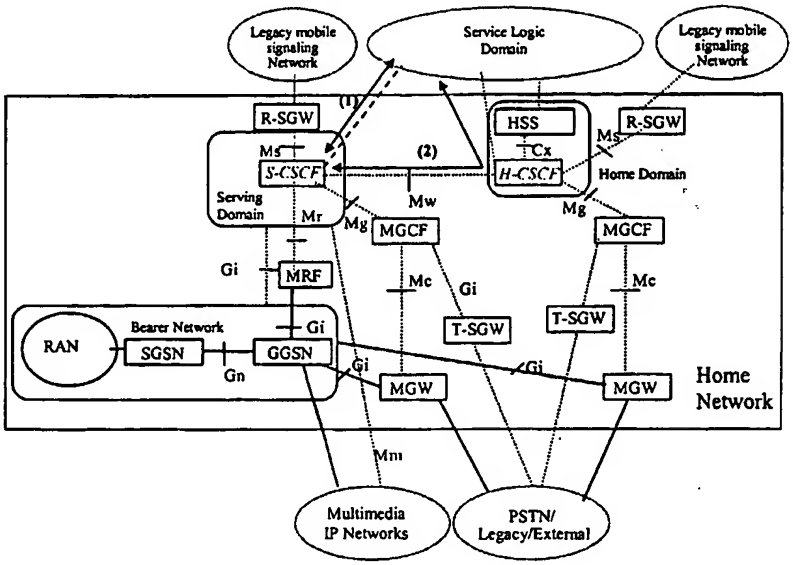


Figure 9-2: Scenario 1 applied to roaming inside a single network

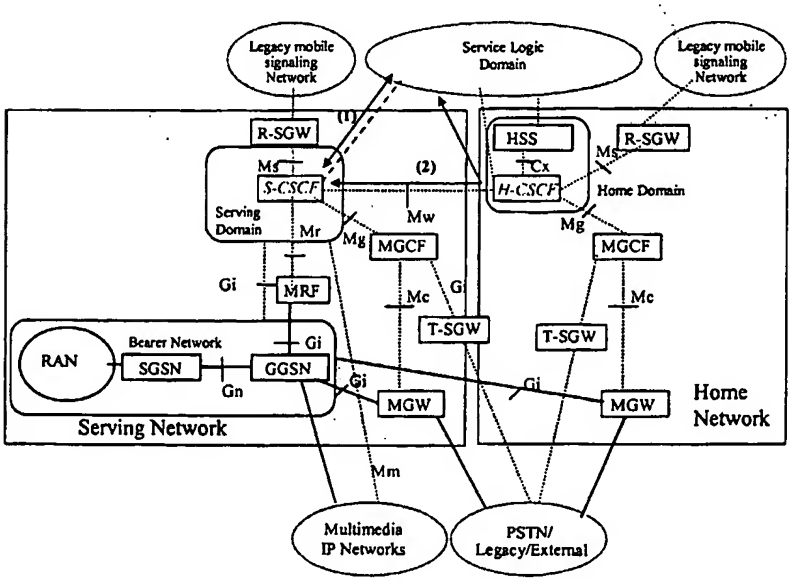


Figure 9-3: Scenario 1 applied to roaming between networks

The following points characterise scenario 1:

- both a Home CSCF and Serving CSCF are present and active;
- MT calls are routed to the Serving CSCF through the Home CSCF;
- non-basic services (e.g. supplementary services, etc.) invoked by MO calls and requiring interaction with service logic specific to the home network operator can be provided in two ways:
- Serving CSCF has a direct interface to the service logic (e.g. direct interface to a SCP in case of IN), case (1) in the above pictures;
- only Home CSCF has access to the service logic and the two CSCFs co-operate in order to provide the service, case (2) in the above pictures;
- MT Calls which reach the Serving CSCF are handled by the Servering CSCF for basic voice services;
- MT calls invoking non-basic services are handled as described for MO calls;
- user plane for MT calls is routed from the originating network to the serving network through the home network (i.e. from PSTN to MGW to GGSN in the serving network);
- user plane for MO calls to non Release 2000 all IP networks (and for non-optimised Release 2000 all IP networks to Release 2000 all IP networks calls) is routed from the serving network to PSTN through a MGW in the serving network, thus optimising the routing of user plane. In case of roaming within the same network, MGW can be chosen "close" to the bearer network again in order to optimise the routing of user plane;
- Home CSCF implements Incoming Call Screening triggers (i.e. triggers for supplementary services and IN services for incoming calls) and relays the call control signalling to the Serving CSCF address retrieved during the HSS interrogation.

9.3.3 Scenario 2

The following pictures show respectively the roaming scenario 2 applied to roaming inside a single network and applied to roaming between networks.

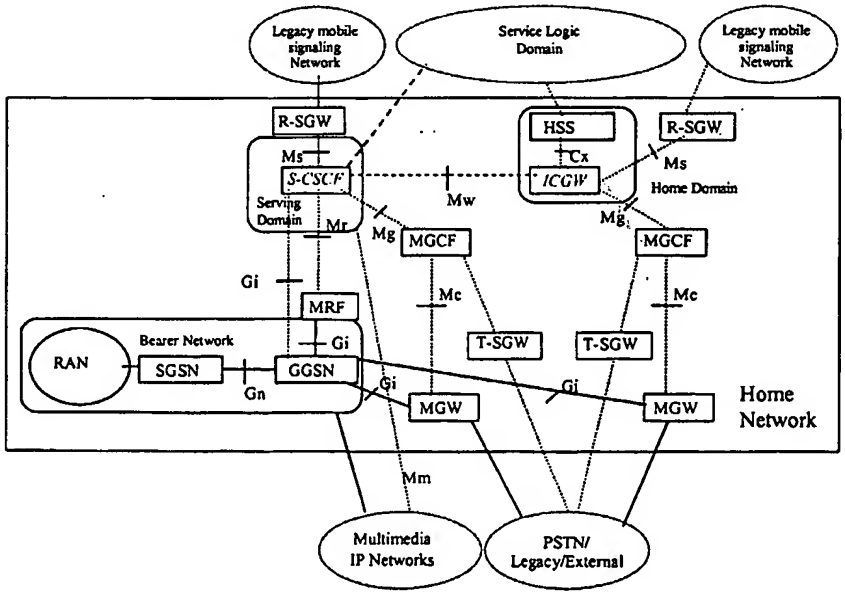


Figure 9-4: Scenario 2 applied to roaming inside a single network

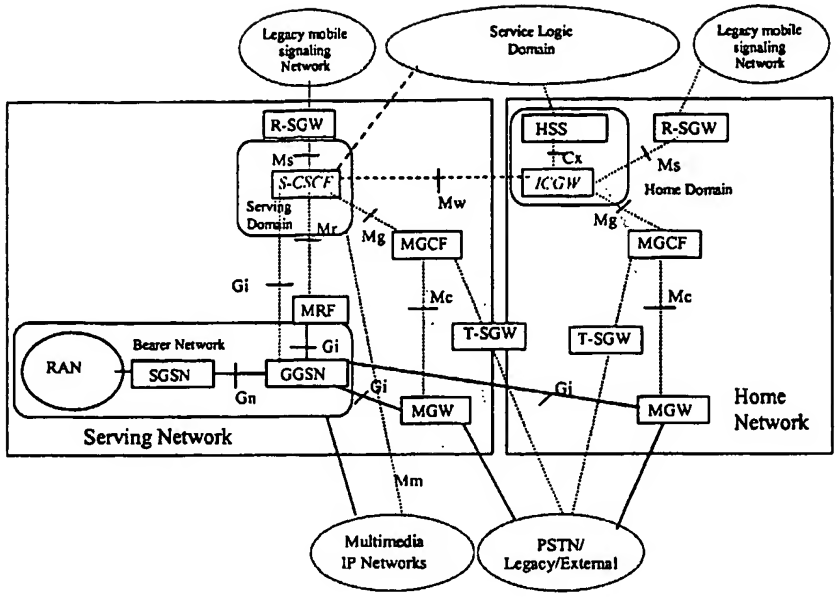


Figure 9-5: Scenario 2 applied to roaming between networks

The following points characterise scenario 2:

- only a Serving CSCF is present;
- MT calls are routed to the Serving CSCF through the ICGW in the Home Domain/Network;
- all the services (basic and non-basic) invoked during MT and MO calls are provided by Serving CSCF;
- If "non basic" service means non standardised, these services will involve serving CSCF and elements within the Home domain and service logic domain.]
- user plane for MT calls is routed from the originating network to the serving network through the home network (i.e. from PSTN to MGW to GGSN in the serving network);
- user plane for MO calls to non Release 2000 all IP networks (and for non-optimised Release 2000 all IP network to Release 2000 all IP network calls) is routed from the serving network to PSTN through a MGW in the serving network, thus optimising the routing of user plane. In case of roaming within the same network, MGW can be chosen "close" to the bearer network again in order to optimise the routing of user plane.

9.3.4 Scenario 1: Information Flows for Validation

In order to validate scenario 1 proposed above, information flows for registration, location management and call delivery/origination are provided in this section.

The information flows presented in the Call Control and Roaming proposal do not provide many details. Generic names have been chosen for signalling messages. Any resemblance to known existing protocols is due to an attempt simply to proposing information flows immediately comprehensible. Only a restricted subset of the possible information flows is included in the document.

9.3.4.1 Registration and Location Management

- In this version of the Technical Report, only a basic registration procedure is considered.
- The basic registration procedure is composed of three steps:
- GPRS attach: is a plain GPRS attach procedure;

- PDP context activation: a PDP context is set up to support application level signalling;
- application level registration with CSCF.

The latter is considered in the registration flows reported in the follow.

Two basic cases of registration are shown here, in order to provide an initial picture of the application level registration.

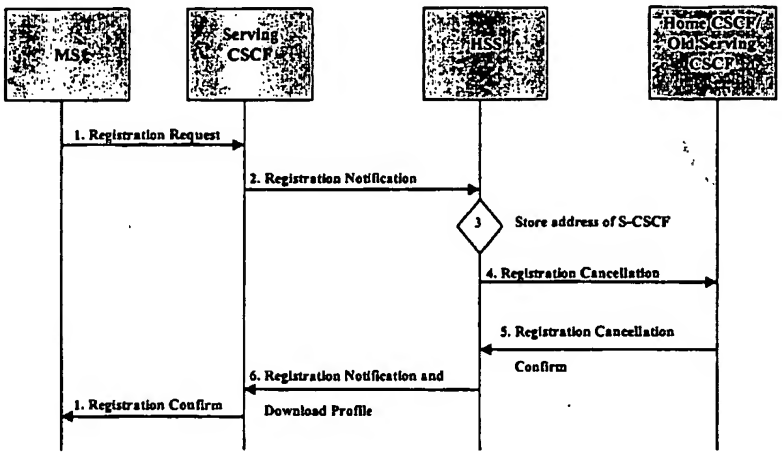


Figure 9-6: Release 2000 all IP network user registering in Release 2000 all IP network Serving Domain

Steps 4 and 5 are optional and take place only in two cases:

- if the user was previously registered in a different Serving CSCF;
- if the user was not previously registered but, during a MT call, HSS determined that a service profile was needed in the Home CSCF to handle the supplementary services triggered by the incoming call (e.g. a forwarded-to multimedia call triggered by the MT call).

Other registration and location management scenarios are FFS.

9.3.4.2 MT/MO Calls

Two call information flows are presented in this version of the Technical Report. The call flows are based on the following call delivery model:

- a call from PSTN towards a DN corresponding to the user is received by one of the MGCF of the Home Network, ISP or corporate LAN domain in the IP multimedia network;
- MGCF reaches the Home CSCF translating the DN;
- Home CSCF interrogates the HSS to retrieve information regarding the user corresponding to DN;
- Home CSCF receives information on how the call has to be routed: in the flows shown here, HSS returns the signalling address of a Serving CSCF, but in general could be the transport address of the MS if there is no

Serving CSCF or a forwarde-to number). Also, HSS might return service profile information in case the user is not registred.

- Home CSCF forwards call signalling to the retrieved address

Regarding user-plane, packets are routed from the MGW in the corporate LAN domain or home network to the GGSN in the bearer network where the user is presently located.

No optimisation has been considered for user-plane of MO calls. In case MO routing optimisation is desired, the MGCF used to route the call towards PSTN can be chosen using different possible criteria.

9.3.4.2.1 Incoming call from PSTN to a Release 2000 all IP network

The following flow describes the call delivery for a MT call from PSTN to a Release 2000 all IP network user addressed through a DN.

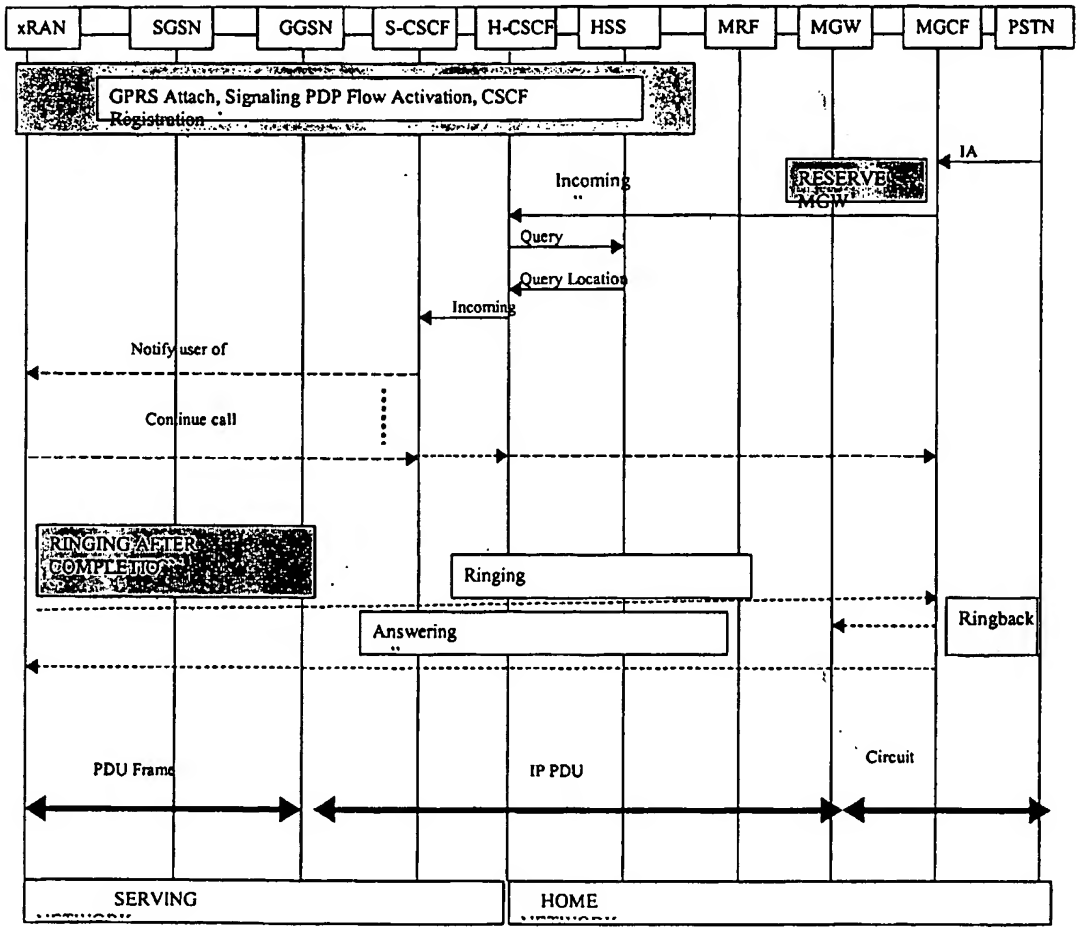


Figure 9-7: Incoming call from PSTN to a Release 2000 all IP network

Issues such as QoS negotiation, policy management, etc. are FFS.

9.3.4.2.2 Call from an 3GPP IP based network/Multimedia IP Network to 3GPP IP network

The following flow assumes that a Release 2000 all IP network user has roamed into a visited network. The flow describes a call from a different Release 2000 all IP network terminated into the home domain of the called user. It is

assumed that the Release 2000 all IP network where the call is coming from, is aware of the addressing plan information (IP based addressing) that allows the call to directly terminate into the H-CSCF (otherwise, it may have been routed through the PSTN and terminated into MGCF).

Details will need to be worked on.

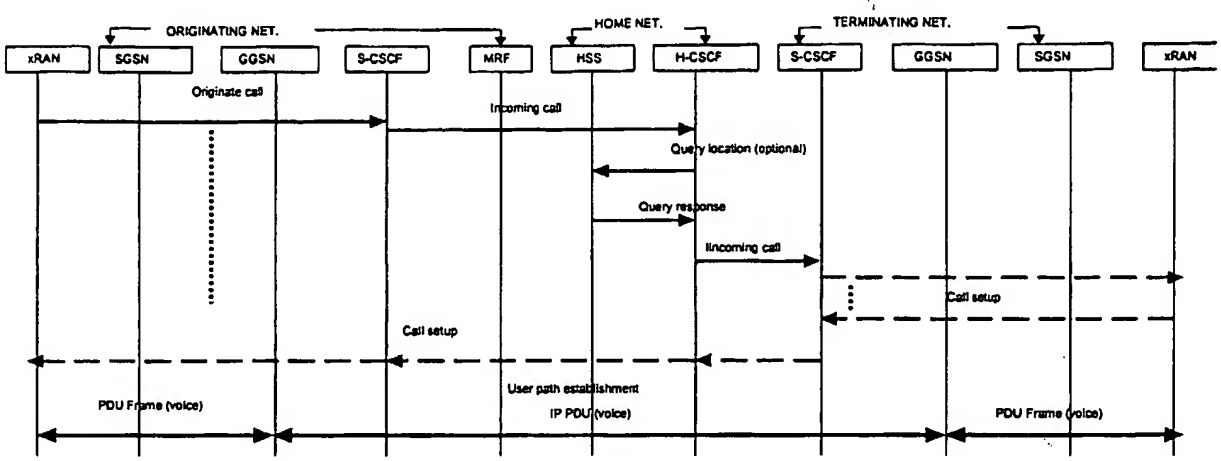


Figure 9-8: Call from Release 2000 all IP network to Release 2000 all IP network

9.3.5 Scenario 2: Information Flows for Validation

No flow will be shown for this version of the document.

9.4 Roaming to Other Networks

In order to ensure compatibility and easy roaming between 2G GSM/GPRS, UMTS R99 and UMTS R00 CS and GPRS domain (excluding the VoIP/multimedia domain), the same mobility procedures are used within and between the 3 kind of networks (storage of the current location in the HSS, use of MAP to update the HSS with the current subscriber location of an user and to download / update the subscriber data in the visited node).

Some enhancements with regard to current mobility procedures are obviously needed:

- in case of a R00 terminal roaming in a MSC/SGSN of a 2G / R99 network:
  - R-SGW relays all the MAP / CAP messages between the HSS / SCP and the functions (MSC, SGSN, ...) handling Call / Session in the 2G / R99 CN.
  - the R00 HSS sends to the MSC/SGSN a MAP_R99 translation of the subscriber data compatible with the data a R99 MSC/SGSN can handle. This should be classical for MAP application context handling.
  - in case of a R00 terminal roaming in R99 network and requesting Multimedia service: as in R99 there is no **standard** way to have a visited GK,
    - Either, in order to get a customized service, the R00 terminal requests service from its Home CSCF in its R00 network. The R-SGW is not impacted in this case.
    - Or, the service of a GK in the visited PLMN is used and the service cannot be customized. The R-SGW is not impacted in this case.

- in case of a R99 terminal roaming in a MSC/SGSN of a R00 network:
  - R-SGW relays all the MAP / CAP messages between the 2G / R99 HLR / SCP and the functions (SGSN, ...) handling Call / Session in the All IP R00 CN. The R00 VLR (in SGSN and/or CS domain) or Call / Session handling function is able to interpret MAP_R99 / CAP_R99 received from R99 HLR / SCP

- in case of a R99 terminal roaming in R00 network and requesting Multimedia service: as in R99 there is no **standard** way to have a visited GK,
  - Either, in order to get a customized service, the R99 terminal requests service from its Home CSCF in its R99 network. The R-SGW is not impacted in this case.
  - Or, the service of a GK in the visited PLMN is used and the service cannot be customized. The R-SGW is not impacted in this case.

There are currently two proposed solutions for roaming. These solutions are not completely in contradiction with each other but actually include a good set of commonalities. However, further work is needed to identify the commonalities and to consolidate the proposals. The two proposals are summarised in this section with references to more detailed descriptions. The presented roaming cases and key issues addressed in the contributions should be further considered and taken as a baseline for further work on roaming model definition for R00 networks. Alternative solutions for roaming are for further study.

9.4.1 Roaming Procedures for R00 networks

One possible solution for the support of roaming in R00 networks is described in Tdoc S2k99117. The contribution covers both roaming between R00 networks and roaming to mobile legacy networks. The contribution only covers the PS-only architecture in UMTS R00. Roaming from UMTS R99 considered in this document is in terms of USIM roaming. The contribution makes certain assumptions on mobility management and network identities that have to be considered as part of the future work. A basic registration procedure is introduced in order to allow the discussion on roaming.

The roaming scenarios described are:

- Roaming within R00 PS domain networks;
- Roaming from R00 PS domain networks to 2G/UMTSR99 networks;
- Roaming from 2G/UMTS R99 networks to R00 PS domain networks;

9.4.2 Overlaid solution to roaming

One roaming solution is to introduce an overlaid personal number service, which keeps track of users registrations (attached to 2G/3G CS and/or PS MultiMedia) and call reception preferences. This enables inter-service as well as inter network roaming for Telephony as classical TeleService Speech in 2G/3G networks and Telephony as the voice component of a MultiMedia service.

The proposed roaming solution is further detailed in Tdoc S2K-99070. The Tdoc elaborates on the driving forces for overlaid roaming and includes examples of user registration and call reception cases.

9.5 Open Issues

The following issues need to be discussed and solved through interaction with the other working groups in Release 2000 all IP network and might require discussion in the plenary.

- Support of multiparty voice and data communications sessions (including the capability for the user or service logic to dynamically add or delete users from an active communications session). The impact on the control plane (e.g. CSCF) and on the user plane (e.g. use of MRF vs. IP multicast) has to be studied.
- UMS structure and functionality has to be defined. As an example, UMS might have additonal interface (e.g. GRIC CSP) to the clean houses as defined in H.225. GRIC Communications, Inc. develops a global intelligent transaction platform (GRIC CSP) that allows diverse networks to interoperate. This platform allows ISPs, Telcos, and emerging carriers to offer multiple IP-based services such as IP Telephony, E-commerce, Internet Roaming, and Internet Faxing. Leveraging its GRIC Alliance Network, a worldwide membership of over 450 major ISPs and telcos in more than 140 countries, GRIC has aggregated an addressable user base of 30 million dial-up users and an estimated 40 million corporate users. For more information on GRIC see http://www.gric.org/ Routing optimisation for calls originated in a Release 2000 all IP network towards a MS of

another Release 2000 all IP network and addressed using a DN (Directory Number) in order to bypass the PSTN has to be discussed.

- Also, routing of MT calls towards LN has to be specified.
- CSCF discovery has to be discussed and solutions have to be presented at the next meetings.
- Impact of requirements regarding backward compatibility with 2G networks (e.g. RAN, terminals, call control and roaming, services) needs to be addressed.
- Location confidentiality issues have to be considered versus optimisation of signalling and transport paths.
- The Release 2000 all IP network must provide the capability for service logic to deny or re-route voice and/or data communication requests. This capability has to apply to both incoming and to user initiated communication requests.
- Also, signalling flooding problems and malicious attacks to the network have to be considered.
- The issue could affect the architecture in terms of where more firewall functions will have to be implemented.
- Addressing of multiple PDP flow by means of the same IP address has to be solved as an issue if not yet standardised for GPRS.
- Details regarding the set of vocoders supported in Release 2000 all IP network and the vocoder negotiation mechanism need to be investigated.
- Signalling PDP flow can be activated when the MS attaches/registers with the network and kept alive for all the duration of the attach/registration session (dorment signalling PDP flow), or it can be activated on demand. The choice between the two options has to be discussed considering paging issues, load due to MM for the dorment PDP context even when the MS is idle, and the impact on the MS.
- Is T.120 considered as real-time application, i.e. do T.120 components of multimedia calls have to be controlled by a CSCF?
- Database queries in the HSS and other network databases has to be discussed and defined.
- Addressing of network entities and address translation need to be defined (e.g. It is assumed that MGCF has the ability to retrieve the CSCF/MSC server corresponding to a DN; how MGCF does it is FFS).

- Storage of 2G profiles in Release 2000 all IP network in order to support roaming to 2G legacy networks has to be discussed. In particular, the presence of a 2G HLR functionality in HSS need to be discussed and possible alternatives evaluated.
- Long term issues and assumptions, i.e. not related only to R00, should be considered in order to have a future proof solution. As an example, long terms requirements on addressing mechanisms (e.g. dynamic vs. static, IPv4 vs. IPv6) should be considered.
- a "VLR" type functionality capable of caching the service profile for the user in a serving domain needs to be defined. The functionality should not be related to location management but focus on service profile.
- Differences in the definition of the term "Location" as understood within the cellular/wireless community and in the IETF community have not been addressed yet in this document. This needs to be harmonized.
- Provisioning of location-based services (e.g. service based on geo-location information) and the impact on the architecture need to be considered.
- QoS and security issues have not been addressed yet.

10 Service Platform Impacts

10.1 3GPP Release 2000 Service Architecture

This section describes how the 3GPP release 99 service architecture [3] can be applied to the 3GPP release 2000 network by extending the VHE/OSA concept to the Multi-Media core network. This can be done by providing an application interface (as described in VHE specification [3]) from the CSCF, see Figure 10-1.. As VHE expect the service to be located in the home domain of the end-user (the Home Environment), other network elements besides the CSCF may be needed to provide an roaming architecture that allow the serving domain to pass control to the home domain where the service logic resides. [1]

Note: The architectures in Figure 10-1 and Figure 10-2 show the CAMEL Service Environment (CSE) which is not shown in the Reference Architecture in Section 5.

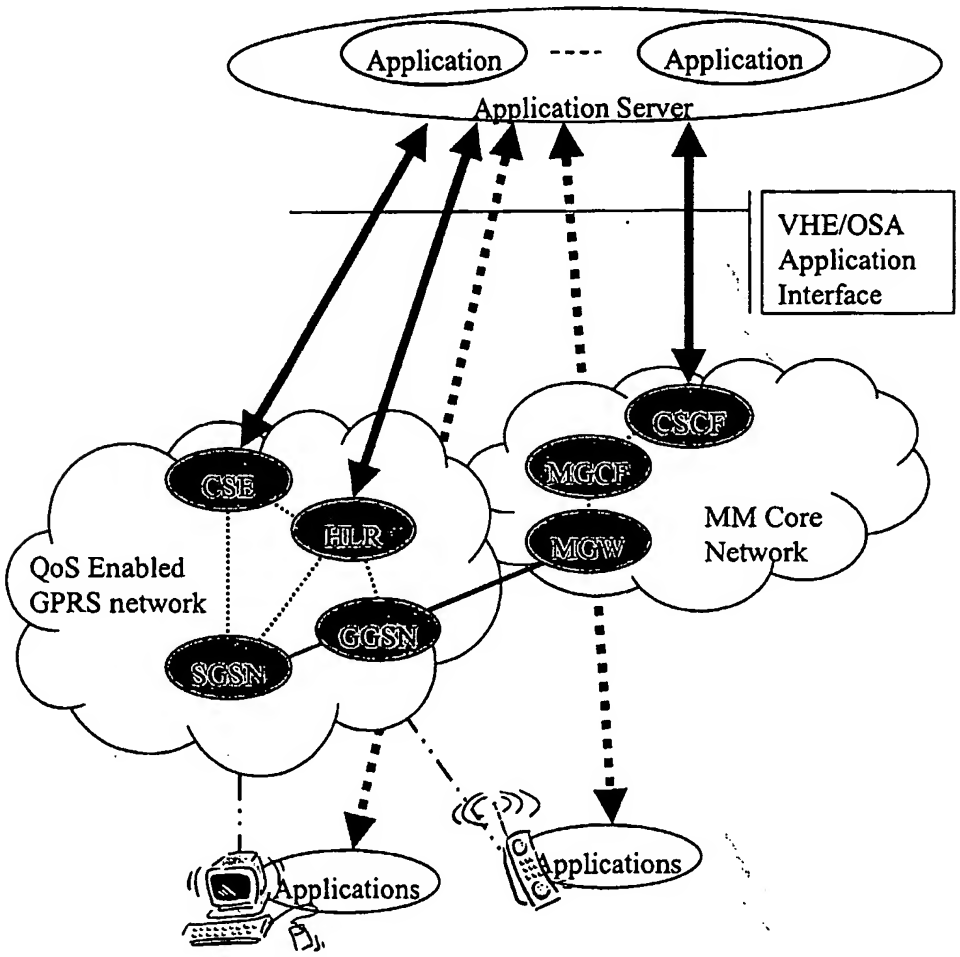


Figure 10-1: 3GPP Release 2000 service architecture

The Open Service Architecture consists of three parts, as illustrated in Figure 10-1(note that the figure is not meant to be exhaustive of all interrelationships):

Within the current 3GPP specifications this is achieved via CAMEL providing the CAP interface between serving and home network.

- **Applications**, e.g. VPN, conferencing, location based applications. These applications are implemented in one or more Application Servers;

- **Framework**, providing the applications with basic services that enable applications to make use of the service capabilities in the network. Examples of framework services are Authentication, Registration and Discovery;

- **Service Capabilities**, providing the applications with services that are abstractions from underlying network functionality. Examples of services offered by the Service Capabilities are Call Control, Message Transfer and Location. Services are possibly provided by more than one Service Capability Server. For example, the Call Control service might be provided by CAMEL and MexE. The Service Capability Servers taken into account for UMTS Release 99 are CAMEL, MExE, SAT and HLR.

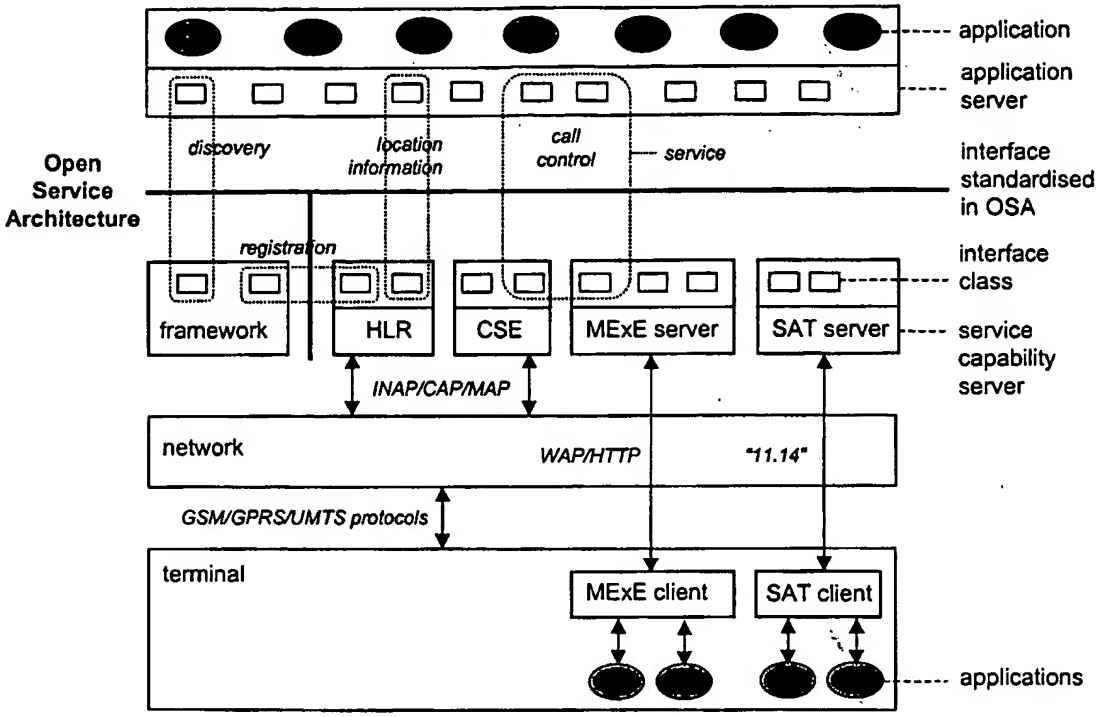


Figure 4

Figure 10-2: Overview of Open Service Architecture

Note: This may not be in line with the latest version of the VHE/OSA Stage 2 document

10.2 IN based Services

The IN based service is one example of legacy services and the IN based service logic is one example of how legacy services may be introduced to the 3GPP Release 2000 networks. This IN based service logic may need to be enhanced in 3GPP Release 2000 networks, based on the proposed architecture, when full support for multi-media is required.

When only voice/audio has to be supported several options exist:

1. Re-route call to legacy system. This is applicable to very specific services such as 800- and 900- services.

2. Provide 'INAP' like interfaces between the 3GPP Release 2000 network functional blocks (e.g. CSCF) and a legacy SCP. These interfaces will be used for inter- and intra-network connections, and as such should be based on a suitable INAP protocol (e.g. CAP).

3. Provide new interfaces between legacy IN and 3GPP Release 2000 network functions, allowing the AS to access the application in the legacy SCP.

It shall be noted that the Operator Specific Services defined for the QoS enabled GPRS network are still available and apply for the access bearers towards the MM core network. This will for instance enable the pre-paid charging of the GPRS bearer.

As indicated in the sections below, option 2 (section 10.2.1) will allow the possibility that existing services can be upgraded to provide 3G users with a seamless transition of familiar 2G services, especially whilst roaming. Option 3 (section 10.2.2) does not have this advantage but benefits from future proofing. It would thus seem appropriate to allow the two options to co-exist, where support of legacy CAMEL services would be carried via option 2 and enhanced/future services can be provided via OSA principles as outlined via option 3 or a combination of both.

10.2.1 'INAP' based interface between legacy SCP and R00 network entities

In this option, the classical IN model is extended to include the CSCF as a node capable of supporting a service switching function (based on the 23.078 specification of the gsmSSF) and a transaction based protocol (TCAP) with CAP. Conceptually, a new functional entity is introduced between the CSCF and the CSE. This functional entity, called a *softSSF*, can potentially be based on a modified release 99 GMSC, VMSC and VLR functionality, where call control, billing and database functions are retained or enhanced. This softSSF interacts with the CSE via CAP and interacts with the CSCF either via an internal interface (if it is co-located with the CSCF) or via an open interface based on OSA concepts (if the softSSF is deployed as a Service Capability Server). Some changes to CAP can be expected to take into account the impact of the underlying IP call control. The CSE is able to offer its services via defined open interfaces based on OSA principles. These services can be implemented by the CSE via the CAP interfaces to the SGSN and the CSCF. (Note the CSCF may also offer its services via defined open interface based on OSA principles).

This option benefits for the extensive re-use of standardised process (gsmSSF), protocol (CAP) and already deployed services.

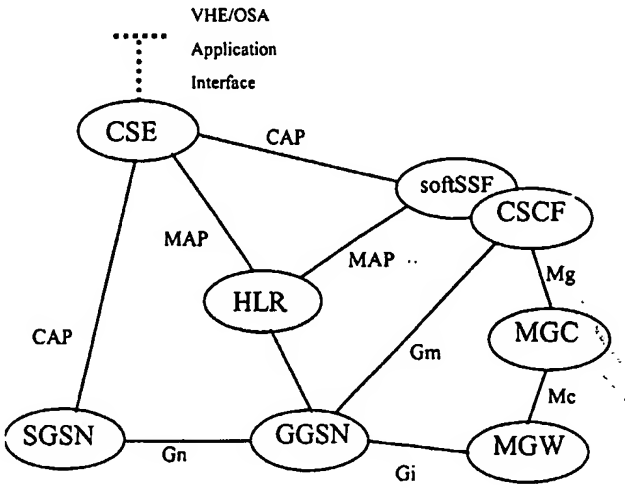


Figure 10-2 Functional Architecture to support option 2

10.2.1.1 Advantages

- Maximises the re-use of existing functional entities, protocols and services. Such reuse decreases the development and ownership costs allowing existing familiar 2G services to be provided to 3G users from an early stage.
- Minimum changes to the CSE for the support of legacy services. There are several IN/CAMEL services already deployed PrePaid, VPN, Mobile Number Portability, etc which can be used in a voice over IP network.
- Potentially, CAP/TCAP can be carrier over IP (this requires further study/contributions)

- This approach is in line with the work currently underway in ETSI SPAN 3 (Services and Protocols), in particular a work item addressing IN support for voice over IP on the H.323 architecture and associated protocols in association with the TIPHON project. The study will investigate how an H.323 gatekeeper can act as a virtual Service Switching Point (SSP). It is worth noting that ETSI plan to harmonise the fixed line IN protocol (ETSI Core INAP CS3.1) and mobile equivalent (CAP Phase 3) into a common protocol targeted for ETSI Core INAP CS4.

10.2.1.2 Disadvantages

- Introduces new functional entity 'softSSF', which provides the necessary mapping between the CSCF and the CSE. However, this functional entity is based on the functions already provided by a VMSC/GMSC, where already standardised process such as the gsmSSF can be reused. The interface between the CSCF and the softSSF requires further study.

10.2.2 New open interface between legacy SCP and R00 network entities

This option adopts the OSA principles where the service capability servers such as the CSCF and the legacy SCF have defined APIs that allow applications in separate application servers to use the features offered by these SCS. This approach allows the service features that are provided by the SCS to be made available to applications designers with out having detailed knowledge of the specific protocols. Currently, the SCSs reflect the service capabilities in UMTS phase1, and CAMEL is one example. From TS 22.121 - "A service capability server consists of one or several server components. Taking CAMEL Services as an example, the server components could be Call Control, Location/Positioning, PLMN Information & Notifications. Each of these server components offers its services via defined open interfaces, and implements these by using GSM/UMTS protocols".

The problem is that within an all IP network, the above mentioned server components are not all available via the CSE as there is no underlying protocol between the CSE and the network for call control (apart form the CSE/SGSN interface). To provide the same functionality provided by existing legacy services, new applications will have to be created that make some limited use of the features that may be offered by the CSE (for example database lookup), plus service features offered in the CSCF.

APIs supporting client-server models will exist. These APIs will enable the users to access service logic via the UE and specialist servers (e.g. MexE). The interface these specialist servers and the CSCF is for further study.

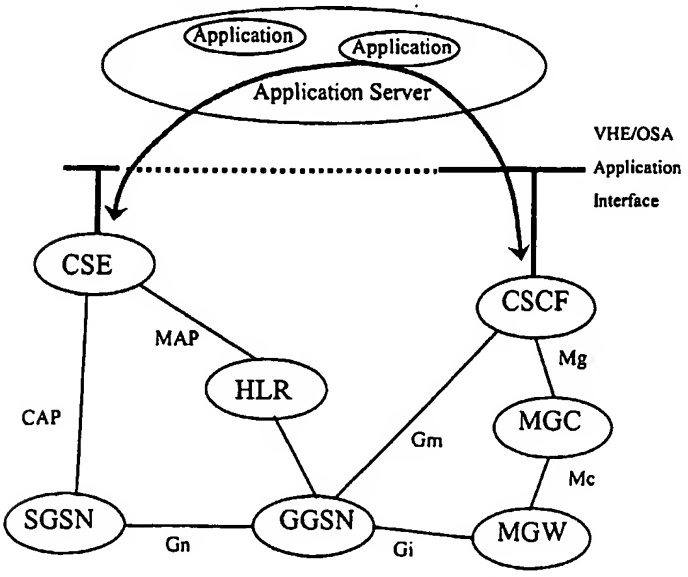


Figure 10-3 Functional Architecture for option 3

10.2.2.1 Advantages

- An open interface based on OSA concepts. APIs that may interface with the CSCF and the CSE are expected to become available allowing specific protocols to be hidden from the service/application designers
- Easier deployment of new/enhanced services for multimedia applications.

10.2.2.2 Disadvantages

- When considering existing CSE based services, little re-use is made of already standardised protocols, services and processes. More significantly, new processes and protocols must be re-defined and re-implemented for services that already exists, increasing development and ownership costs.
- Requires new applications on an application server to be created in order to support legacy services.
- Impact on the legacy CSE and services considered greater than option 2.

10.3 Issues requiring further contributions

The following issues require further contributions:

- Applications may reside not only in Application Servers (AS) but also in terminals.
- Options for sharing applications or parts of them between AS and terminals
- Which elements, beside the CSCF, will provide API for application design (aligned with VHE/OSA)
- Terminal shall also provide API for application design (aligned with MExE)
- Which new Service Capabilities/Service Capability Features are needed for 3GPP Release 2000 (e.g. WIN)
- Specific implementation cases of the proposed architecture should be provided.
- If and how 3GPP Release 2000 service features could be made accessible to 2G terminals via 2G networks
- If and how 3GPP Release 2000 service features could be made accessible to dual mode 2G/3G terminals via 2G networks

11 Security

There will be a common authentication scheme for the terminals operating in the all-IP mode, which will be SIM/USIM based. It is required that all-IP terminals will be able to register and provide basic service when used with a 3GPP SIM/USIM.

12 Work Plan

12.1 Milestones for Release 00

3GPP has the objective of producing the second release of specification for UMTS by the end of 2000. The project management for this work will need to include the elements of work package definition, the interdependency of these work packages and their scheduling. As the work is undertaken in the various TSGs and WGs in 3GPP there will need to be agreement across these groups on the overall plan. Subsequent communication between these groups on such issues as changes to schedules and requirements will be essential. An additional task relating to communication is the

reviewing of requirements documents to identify technical problems in the implementation at an early stage. The completion of release 99 has the potential to impact the availability of resources and hence the time scales for release 00.

Note: "The text that follows is of more general nature than the rest of the document. This has been included to show the framework for release 2000, of which an All IP Architecture may be one Work Item. The dates shown are assumed dates and need to be verified by TSG-S2."

12.1.1 Release 00 milestones

3GPP has not yet agreed overall milestones for release 00. For the purposes development of a high-level work plan the following key milestones are proposed.

July 99	3GPP All-IP network feasibility study started
Sept 99	TSG-S2 R00 Ad Hoc will submit the results of the TSG-S2 for approval
Oct 99	After TSG-S2 approval, TSG-S2 R00 Ad Hoc Group results will be submitted to TSG-S for approval
Oct 99	After TSG-S approval Project planning work for R00 (including all IP option) will be started at TSG-S2 project planning ad hoc groups (These groups have adequate participation from all relevant TSGs/WGs).
Dec 99	After TSG-S2 approval the detailed workplan for R00 (including all-IP network option) will be submitted to TSG-S for approval.
Dec 99	Release 00 service requirements available
Jan 00	TSG-SA2 completes first draft architecture for all-IP network.
Jan 00-Dec 00	Work within 3GPP TSGs and WGs proceed according the TSG-S approved R00 Project Plan
Dec 00	Release 00 specifications completed including all-IP option

A high level PERT chart is given overleaf to form the basis of the planning for releases beyond R99. The reviewed first drafts of the service and architecture specifications should be available before the end of 1999.

12.1.2 Detailed activity plan

Date	Meeting group	Proposed activity
August 23 – 27	S2	Progress architectural study
September 13 - 17	S2	Joint S1 – S2 activity on R00 - Finalize the requirements for the architectural study and identify the key issues. Finalize the proposed R00 architecture at the TSG-S2 R00 ad hoc group.
Late September/early October	S2 R00 Ad Hoc Group meeting (provisional)	Possible refinements to the outcome of the ad hoc group
September 29 - 1	S1	Review input on service requirements for R00
October 11 - 13	SA Plenary	Approve the results of TSG-S2 R00 Ad Hoc Group
October 25 – 29	S2	Start the specification of the architecture, and detailed work plan. Initiate the S2 project coordination adhoc groups work for R00 based on the approved results from TSG-SA Plenary as well as results from other

		groups, e.g. the Mobile IP ad-hoc group.
November 29 - 3	S1	Review service requirement and architecture
November 29 - 3	S2	Develop architecture specifications
December 15 - 17	SA Plenary	Approve detailed work plan from TSG-S2 project coordination adhoc groups

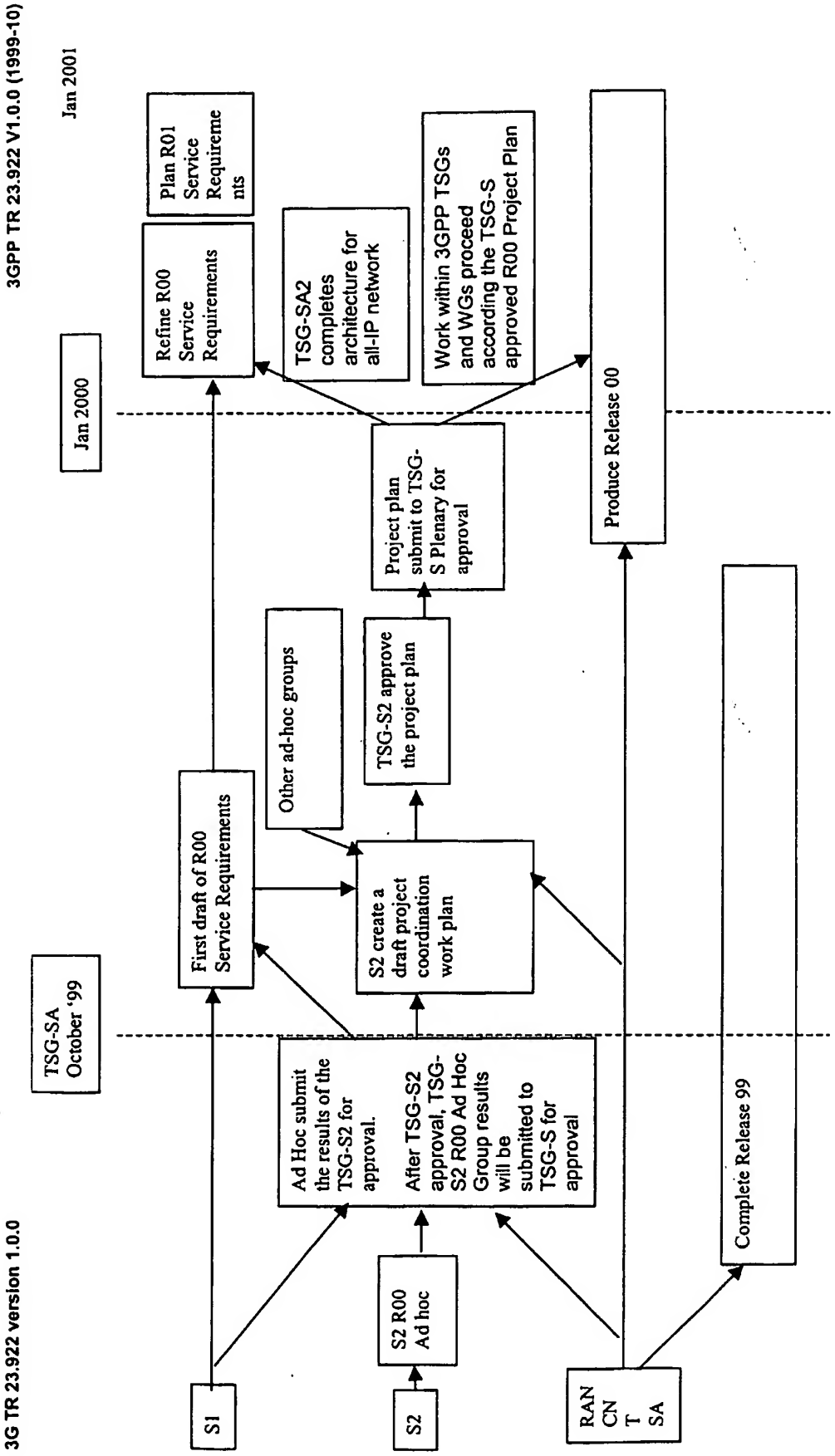


Figure 14-1: 3GPP Standardisation Activities Beyond Release 99

History

Document history		
V0.0.0	July 1999	Creation of document.
V0.0.1	11 Aug 1999	Updated after R2000 Ad Hoc Swindon, 10-11th August. Scope is clarified and new text added to requirements, Service Platform sections and a new architecture sub-section on support for R99 terminals. Architecture changes: new CSCF <-> CSCF interface. MGW split into MGW and Transport signalling gateway. SGW to legacy network labelled as roaming gateway.
V0.0.2	1 Sept 1999	Updated after R00 Ad Hoc, based on tdocs s2k99030, s2k99031, s2k99032, s2k99035, s2k99045, s2k99049
V0.0.3	6th Sept 1999	Updated with new sections reviewed by e-mail. Contains marked changes from v0.0.2
V0.1.0	22 Sept 1999	Updated as per meeting week beg. 13th Sept (Bonn). New text on handover and IP header compression /stripping. Definitions on Mc reference point and on the Media Gateway function al blocks.
V0.1.1	28th Sept 1999	Draft for Ad Hoc Helsinki, changes include addition of HSS in section 5, agreeed solution for support of CS terminals.
V0.1.2	29th Sept 1999	2nd draft for Ad Hoc Helsinki, changes include addressing editor's notes in section 9, addirtion of text to sections on QoS and Security and new subscetions to service Platforms
V0.1.3	30th Sept 1999	Interim.draft at R00 Ad Hoc, Helsinki
V0.1.4	1st Oct 1999	Final draft for approval
V.1.0.0	7th Oct 1999	Prepared for presentation at SA#5. Technical content identical to v.0.1.4

CLAIMS:

1. A packet switched environment, including a functionality of a presence server in an application and services environment.

2. The Packet switched environment of claim 1 wherein the presence server receives a REGISTER message from an interrogating call state control function in the same network.

3. The Packet switched environment of claim 2 wherein the serving call state control function of the presence server network further receives the REGISTER message.

4. The Packet switched environment of claim 2 or claim 3 wherein the presence server receives SUBSCRIBE message from the interrogating call state control function in the same network.

5. The Packet switched environment of claim 4 wherein the interrogating call state control function receives the SUBSCRIBE message from a serving call state control function in a subscribers home network.

6. The Packet switched environment of claim 5 wherein the serving call state control function receives the SUBSCRIBE message from a proxy call state control function in a subscribers home network.

7. The Packet switched environment of claim 6 wherein the proxy call state control function receives the SUBSCRIBE message from the subscriber.

8. The Packet switched environment of any one of claims 2 to 7 wherein the presence server provides a NOTIFY message to the interrogating call state control function of the subscriber's network.

9. The Packet switched environment of claim 8 wherein the interrogating call state control function provides the NOTIFY

message to a proxy call state control function of the subscriber network.

10. The Packet switched environment of claim 9 wherein the proxy call state control function provides the NOTIFY message to the subscriber.

11. The Packet switched environment of claim 1 wherein the presence server receives a plurality of SUBSCRIBE signals from a respective plurality of subscribers in a subscriber network.

12. The Packet switched environment of claim 11 wherein each subscriber provides a SUBSCRIBE message to a proxy call state control function of the subscriber network.

13. The Packet switched environment of claim 12 wherein the proxy call state control function of the subscriber network forwards the respective SUBSCRIBE messages to a serving call state control function of the subscriber network.

14. The Packet switched environment of claim 13 wherein the serving call state control function of the home network provides the respective SUBSCRIBE messages to an interrogating call state control function of the presence server network.

15. The Packet switched environment of claim 14 wherein the interrogating call state control function provides the respective SUBSCRIBE messages to the presence server.

16. The Packet switched environment of any one of claims 11 to 15 wherein the presence server provides a single SUBSCRIBE message to a serving call state control function of the presence server network.

17. The Packet switched environment of claim 16 wherein the presence server receives s single NOTIFY message from the serving call state control function.

18. The Packet switched environment of claim 17 wherein the presence server generates a NOTIFY message for each respective user subscriber.

19. The Packet switched environment of claim 18 wherein the plurality of NOTIFY messages are received by the interrogating call state control function of the subscriber network.

20. The Packet switched environment of claim 19 wherein the interrogating call state control function forwards the NOTIFY messages to a serving call state control function of the subscriber network.

21. The Packet switched environment of claim 20 wherein the serving call state control function forwards the NOTIFY messages to a proxy call state control function of the subscriber network.

22. The Packet switched environment of claim 21 wherein the proxy call state control function forwards the NOTIFY messages to the respective subscribers.

23. The Packet switched environment of claim 2 wherein the presence server receives a REGISTER message from a serving call state control function in the presence server network network.

24. The Packet switched environment of claim 23 wherein the serving call state control function receives the REGISTER message from an interrogating call state control function of the presence server network.

25. The Packet switched environment of claim 24 wherein a subscriber provides a SUBSCRIBE message to a proxy call state control function of the subscriber network.

26. The Packet switched environment of claim 25 wherein the proxy call state control function of the subscriber network

forwards the SUBSCRIBE message to a serving call state control function of the subscriber network.

27. The Packet switched environment of claim 13 wherein the serving call state control function of the subscriber network provides the SUBSCRIBE message to an interrogating call state control function of the presence server network.

28. The Packet switched environment of claim 14 wherein the interrogating call state control function provides the respective SUBSCRIBE messages to a serving call state control function.

29. The Packet switched environment of claim 28 wherein the serving call state control function provides the SUBSCRIBE message to presence server.

30. The Packet switched environment of claim 29 wherein the presence server provides NOTIFY message to the serving call state control function.

31. The Packet switched environment of claim 30 wherein the NOTIFY message is received by the interrogating call state control function of the subscriber network.

32. The Packet switched environment of claim 31 wherein the interrogating call state control function forwards the NOTIFY message to a serving call state control function of the subscriber network.

33. The Packet switched environment of claim 32 wherein the serving call state control function forwards the NOTIFY message to a proxy call state control function of the subscriber network.

34. The Packet switched environment of claim 33 wherein the proxy call state control function forwards the NOTIFY messages to the subscriber.

35. The packet switched environment of any preceding claim, wherein the environment is an internet protocol multimedia environment.

36. The packet switched network of claim 35 wherein the internet protocol multimedia environment is a subsystem of an all-IP telecommunications network.

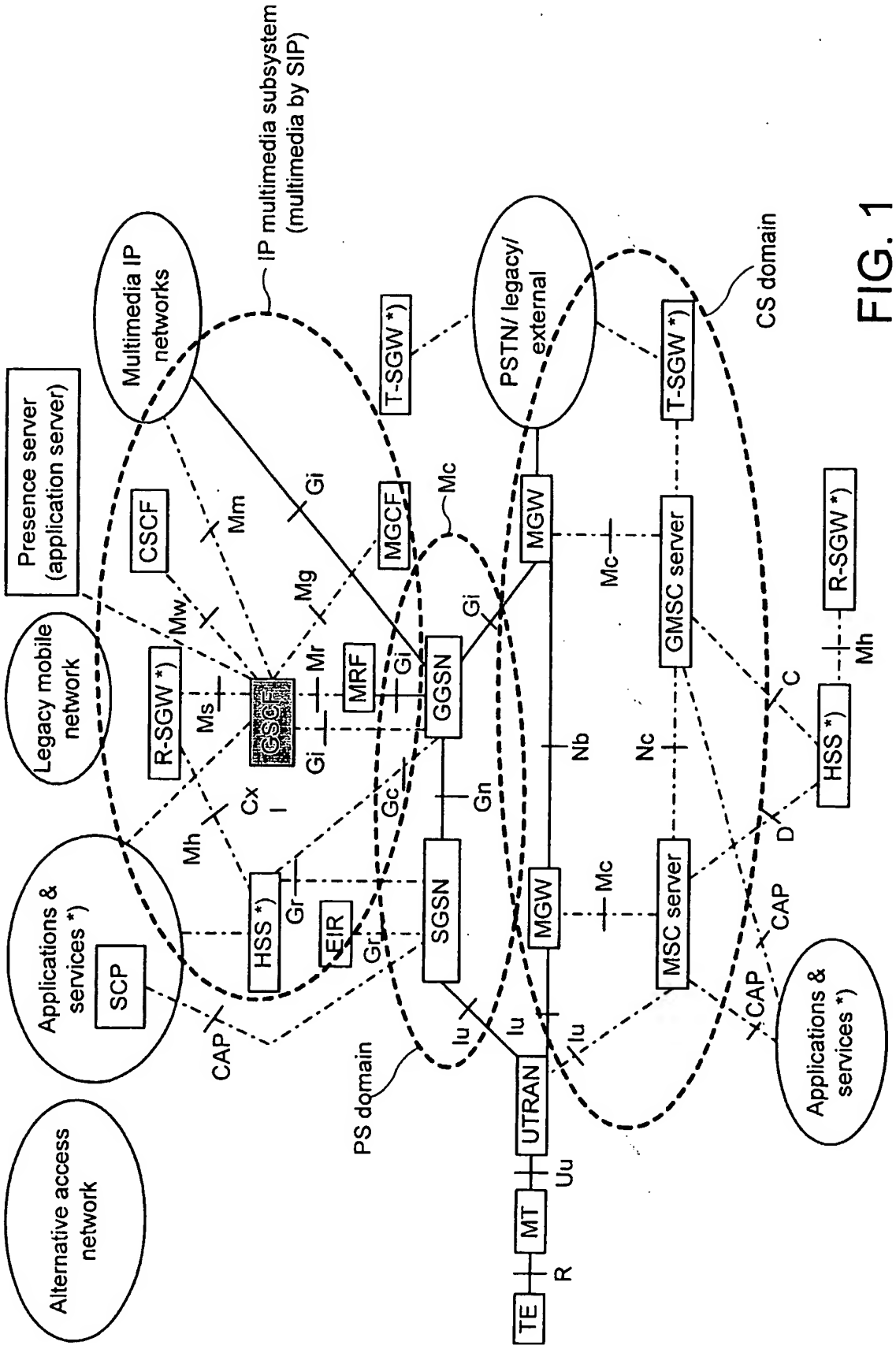

IP multimedia subsystem (multimedia by SIP)
Multimedia IP networks
PSTN/ legacy/ external
CS domain
Presence server (application server)
Legacy mobile network
Applications & services *)
Alternative access network
Applications & services *)
CSCF
R-SGW *)
GSCF
HSS *)
EIR
SCP
MRF
MGCF
T-SGW *)
GGSN
SGSN
MGW
GMSC server
MSC server
R-SGW *)
HSS *)
UTRAN
MT
TE
PS domain
Mm
Mw
Ms
Mg
Mr
Gi
Mc
Cx
Gc
Gr
Gn
Nb
Nc
Mh
C
D
CAP
Iu
Uu
R

FIG. 1

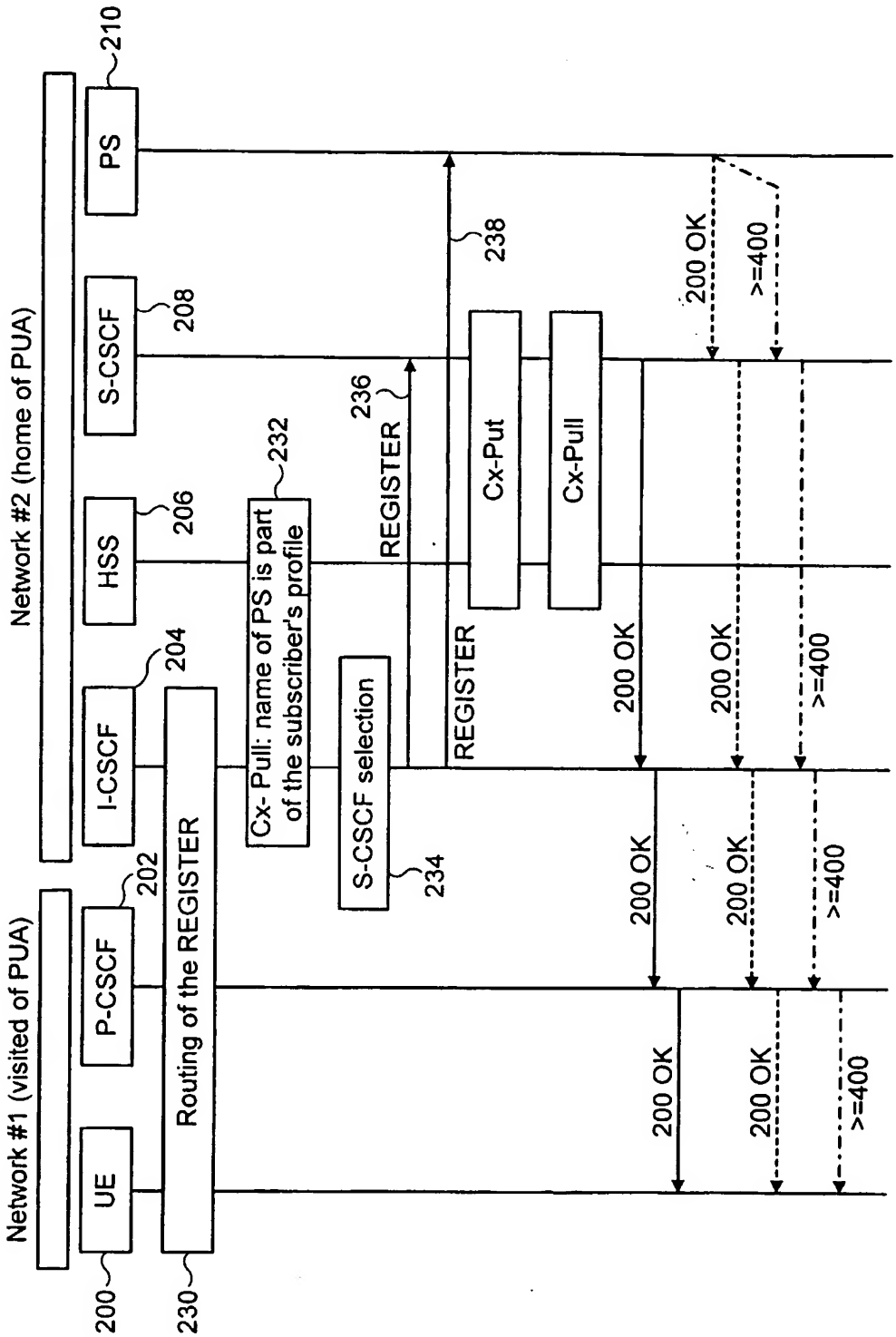

Network #1 (visited of PUA)
Network #2 (home of PUA)
UE
P-CSCF
I-CSCF
HSS
S-CSCF
PS
200
202
204
206
208
210
230
Routing of the REGISTER
Cx- Pull: name of PS is part of the subscriber's profile
232
S-CSCF selection
234
REGISTER
236
REGISTER
238
Cx-Put
Cx-Pull
200 OK
>=400

FIG. 2

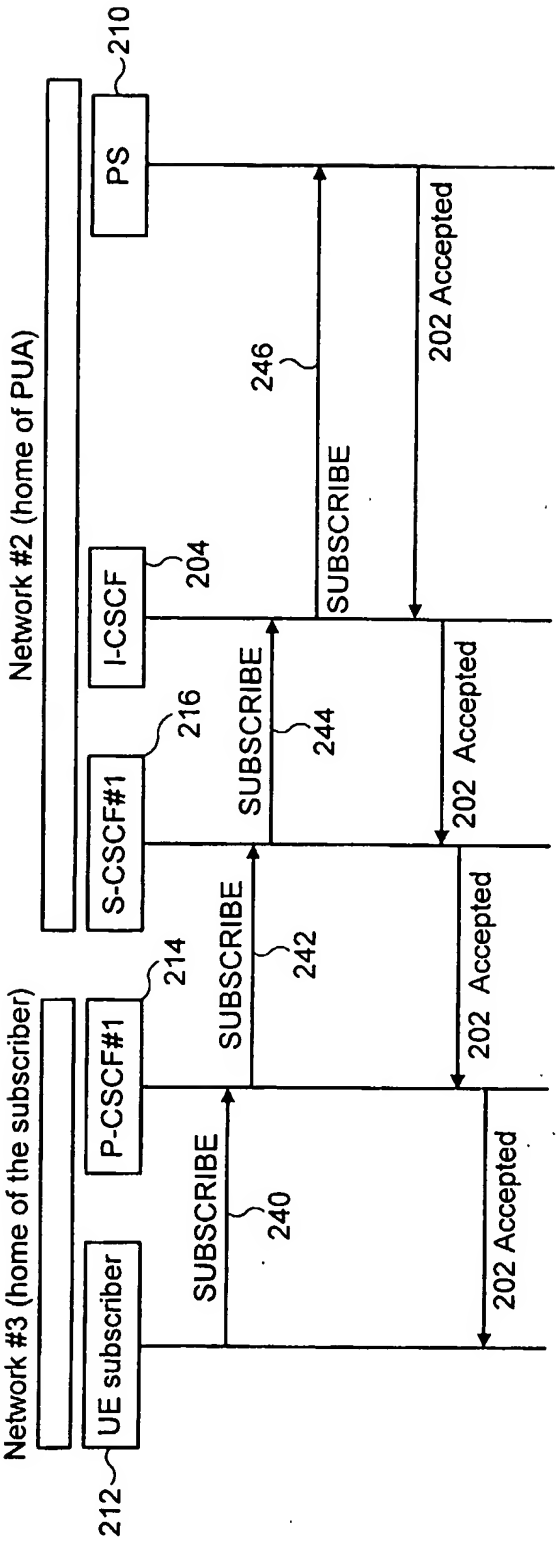

Network #3 (home of the subscriber)
Network #2 (home of PUA)
UE subscriber
212
P-CSCF#1
214
S-CSCF#1
216
I-CSCF
204
PS
210
SUBSCRIBE
240
SUBSCRIBE
242
SUBSCRIBE
244
SUBSCRIBE
246
202 Accepted
202 Accepted
202 Accepted
202 Accepted

FIG. 3

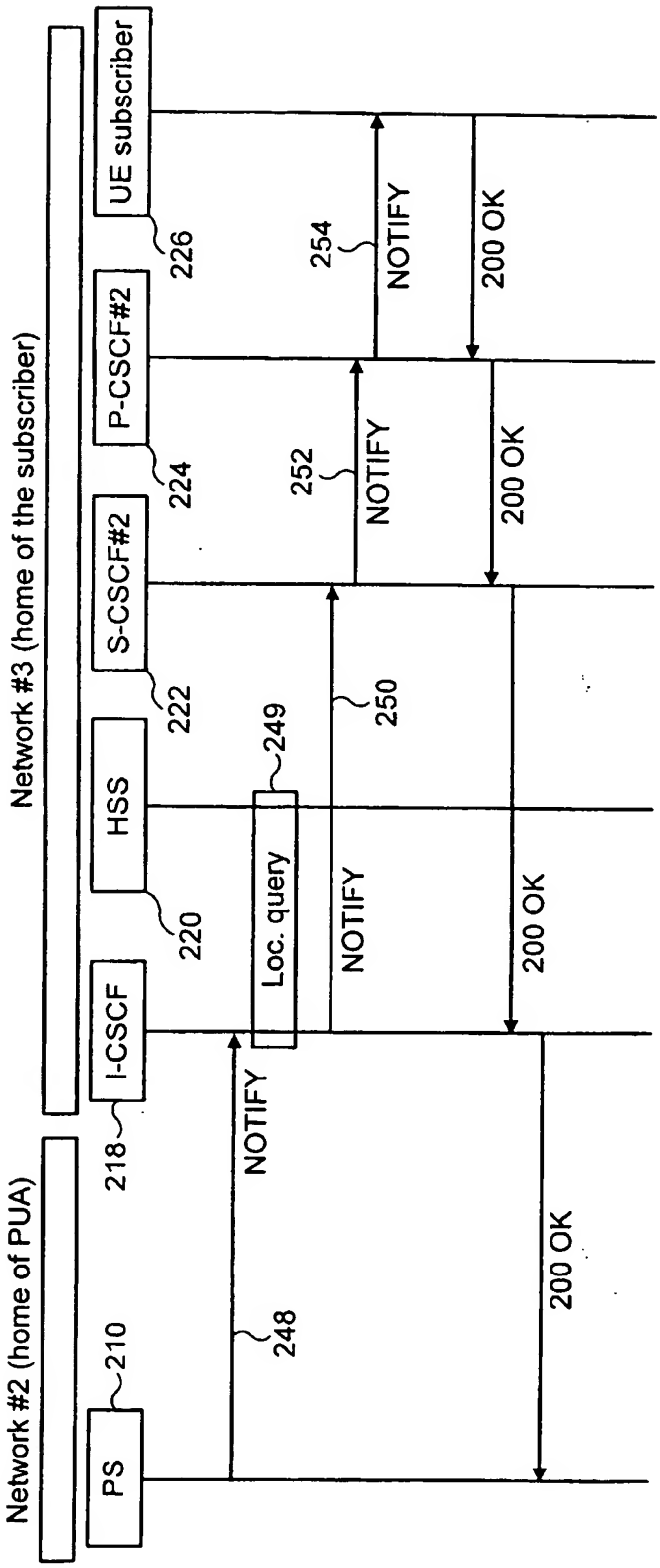

Network #2 (home of PUA)
Network #3 (home of the subscriber)
PS
210
I-CSCF
218
HSS
220
S-CSCF#2
222
P-CSCF#2
224
UE subscriber
226
Loc. query
249
NOTIFY
248
NOTIFY
250
NOTIFY
252
NOTIFY
254
200 OK
200 OK
200 OK
200 OK

FIG. 4

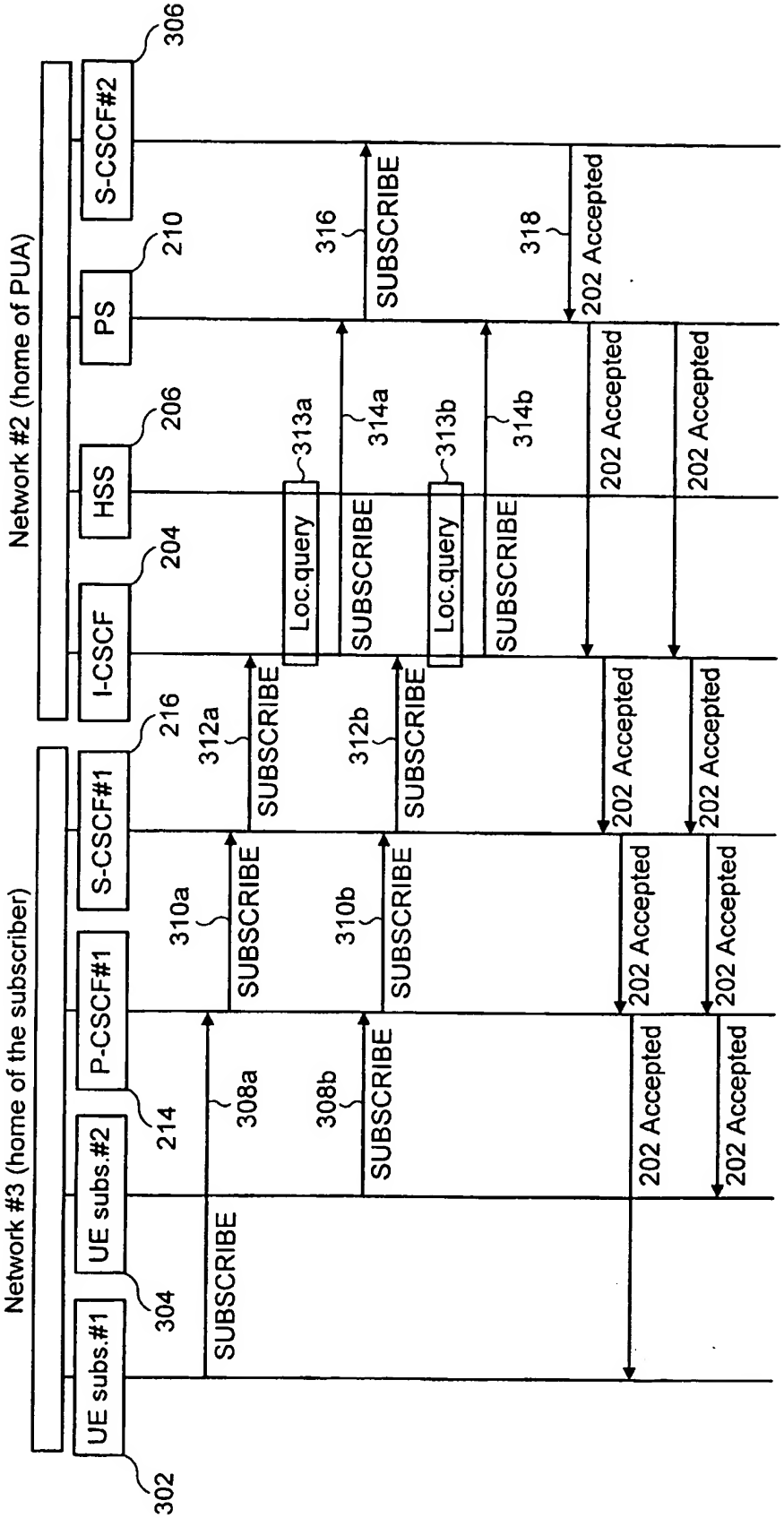

Network #3 (home of the subscriber)
Network #2 (home of PUA)
UE subs.#1
UE subs.#2
P-CSCF#1
S-CSCF#1
I-CSCF
HSS
PS
S-CSCF#2
302
304
214
216
204
206
210
306
SUBSCRIBE
202 Accepted
Loc.query
308a
308b
310a
310b
312a
312b
313a
313b
314a
314b
316
318

FIG. 5

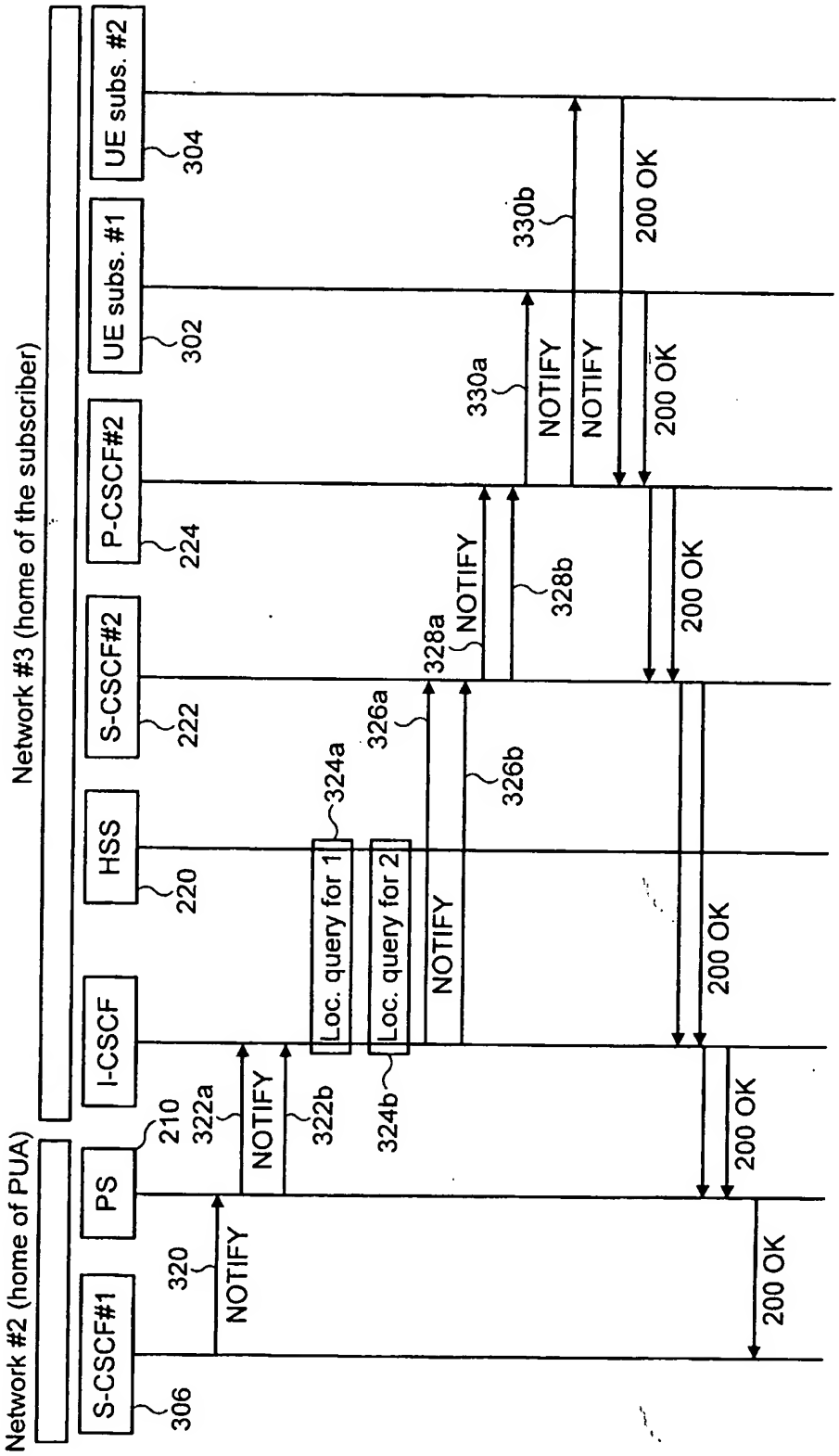

Network #2 (home of PUA)
Network #3 (home of the subscriber)
S-CSCF#1
306
PS
210
I-CSCF
HSS
220
S-CSCF#2
222
P-CSCF#2
224
UE subs. #1
302
UE subs. #2
304
320
NOTIFY
322a
NOTIFY
322b
Loc. query for 1
324a
Loc. query for 2
324b
326a
NOTIFY
326b
328a
NOTIFY
328b
330a
NOTIFY
NOTIFY
330b
200 OK
200 OK
200 OK
200 OK
200 OK
200 OK

FIG. 6

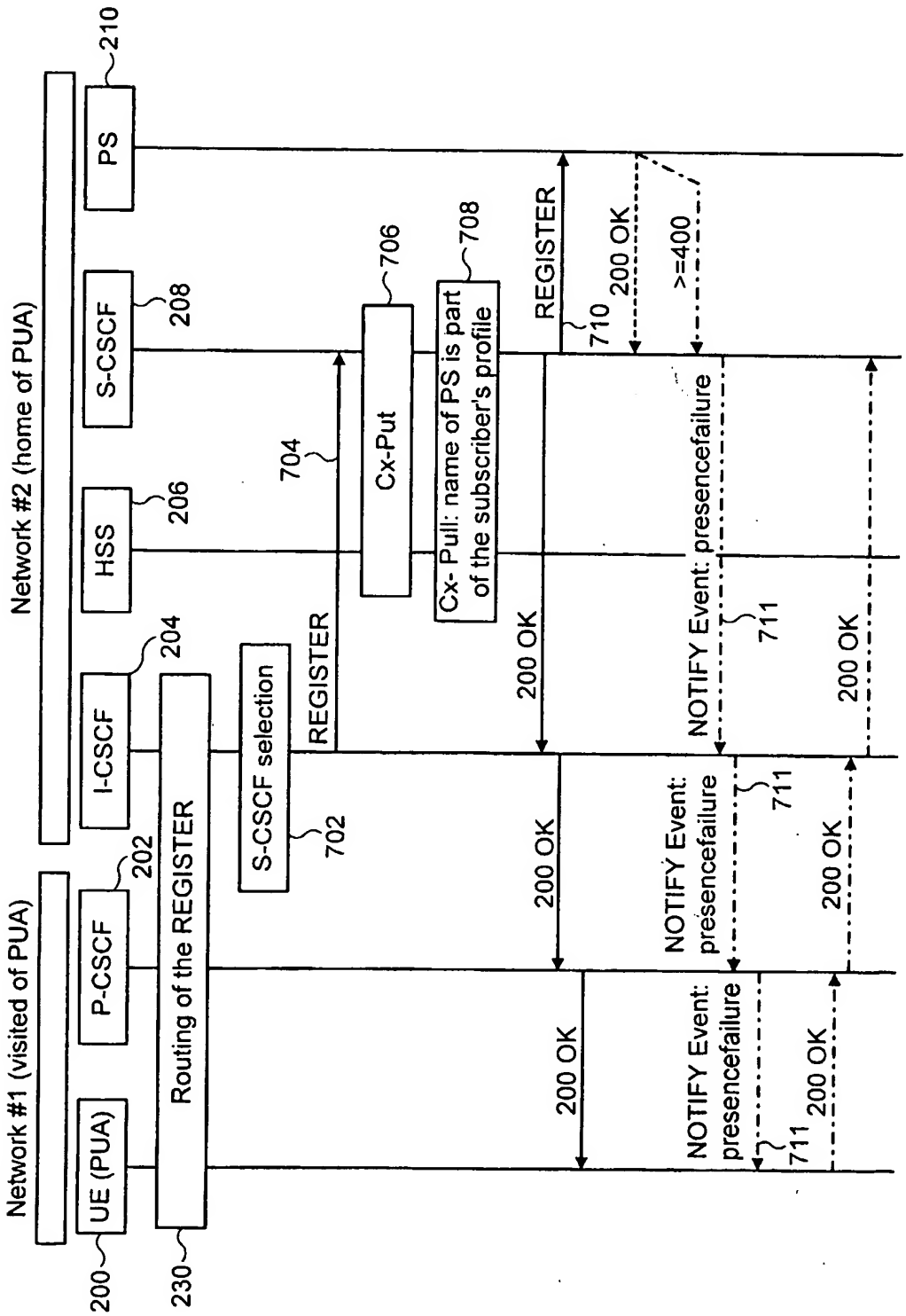

Network #1 (visited of PUA)
Network #2 (home of PUA)
UE (PUA)
P-CSCF
I-CSCF
HSS
S-CSCF
PS
200
202
204
206
208
210
230
Routing of the REGISTER
S-CSCF selection
702
REGISTER
704
Cx-Put
706
Cx- Pull: name of PS is part of the subscriber's profile
708
REGISTER
710
200 OK
>=400
200 OK
NOTIFY Event: presencefailure
711
200 OK
NOTIFY Event: presencefailure
711
200 OK
NOTIFY Event: presencefailure
711
200 OK
200 OK
200 OK

FIG. 7

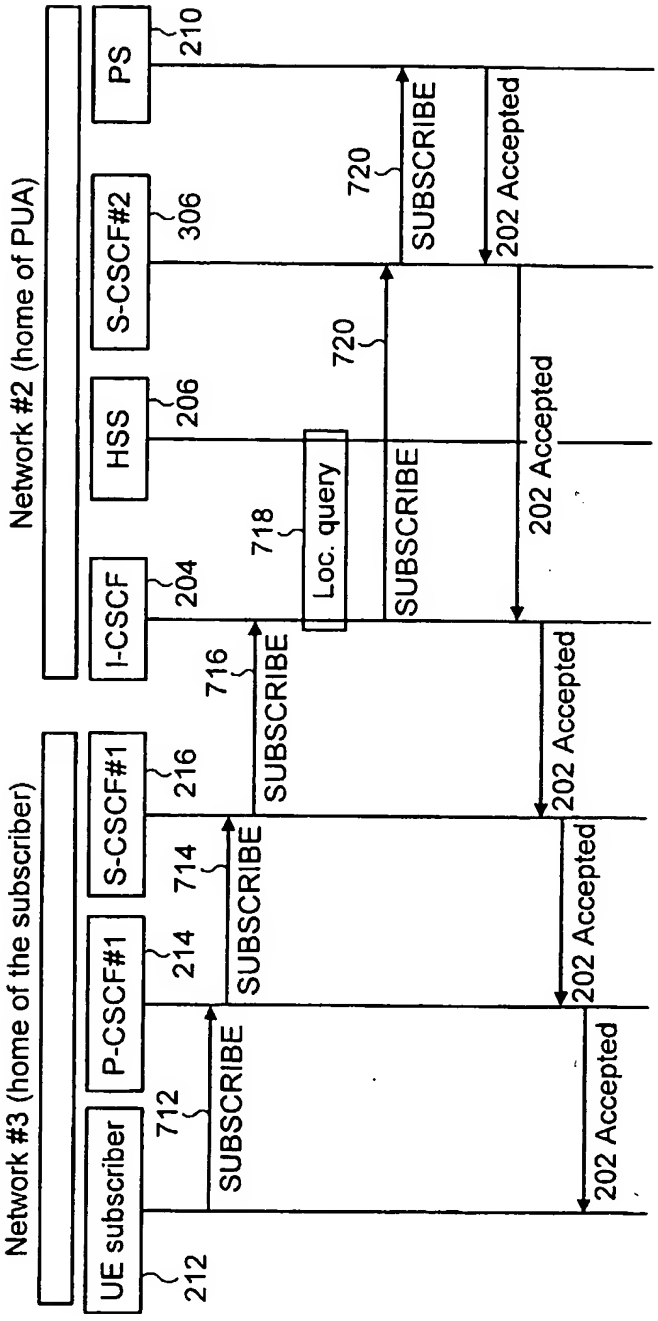

Network #3 (home of the subscriber)
Network #2 (home of PUA)
UE subscriber
212
P-CSCF#1
214
S-CSCF#1
216
I-CSCF
204
HSS
206
S-CSCF#2
306
PS
210
712
SUBSCRIBE
714
SUBSCRIBE
716
SUBSCRIBE
718
Loc. query
720
SUBSCRIBE
720
SUBSCRIBE
202 Accepted
202 Accepted
202 Accepted
202 Accepted
202 Accepted

FIG. 8

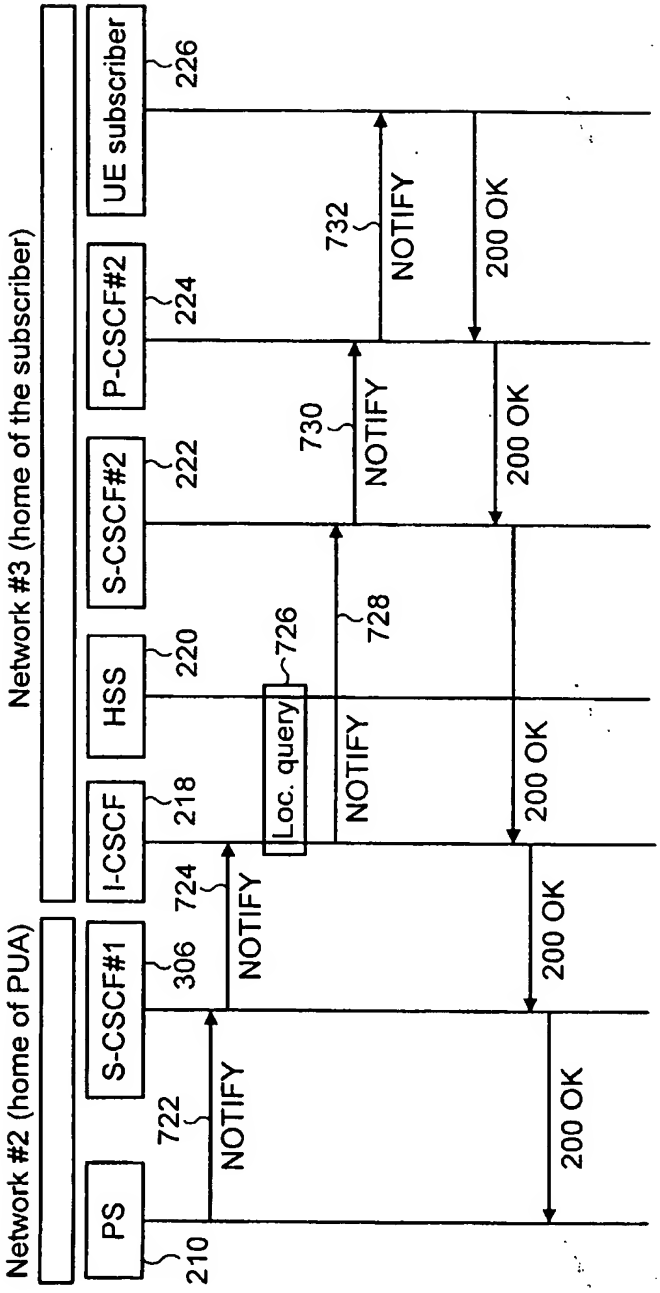

Network #2 (home of PUA)
Network #3 (home of the subscriber)
PS
210
S-CSCF#1
306
I-CSCF
218
HSS
220
S-CSCF#2
222
P-CSCF#2
224
UE subscriber
226
722
NOTIFY
724
NOTIFY
726
Loc. query
728
NOTIFY
730
NOTIFY
732
NOTIFY
200 OK
200 OK
200 OK
200 OK
200 OK

FIG. 9

PATENT COOPERATION TREATY

PCT

DECLARATION OF NON-ESTABLISHMENT OF INTERNATIONAL SEARCH REPORT

(PCT Article 17(2)(a) and Rule 39)

Applicant's or agent's file reference 300327.WO/DJ	**IMPORTANT DECLARATION**	Date of mailing *(day month year)* 13 -09- 2002
International application No. IB02/02212	International filing date *(day month year)* 02.04.2002	(Earliest) Priority Date *(day month year)* 30.03.2001
International Patent Classification (IPC) or both national classification and IPC Ho4Q 3/00, H04Q 7/24		
Applicant Nokia Corporation et al		

This International Searching Authority hereby declares, according to Article 17(2)(a), that **no international search report will be established** on the international application for the reasons indicated below.

1. ☐ The subject matter of the international application relates to:
   - a. ☐ scientific theories.
   - b. ☐ mathematical theories.
   - c. ☐ plant varieties.
   - d. ☐ animal varieties.
   - e. ☐ essentially biological processes for the production of plants and animals, other than microbiological processes and the products of such processes.
   - f. ☐ schemes, rules or methods of doing business.
   - g. ☐ schemes, rules or methods of performing purely mental acts.
   - h. ☐ schemes, rules or methods of playing games.
   - i. ☐ methods for treatment of the human body by surgery or therapy.
   - j. ☐ methods for treatment of the animal body by surgery or therapy.
   - k. ☐ diagnostic methods practised on the human or animal body.
   - l. ☐ mere presentations of information.
   - m. ☐ computer programs for which this International Searching Authority is not equipped to search prior art.
2. ☒ The failure of the following parts of the international application to comply with prescribed requirements prevents a meaningful search from being carried out:
   ☐ the description ☒ the claims ☐ the drawings
3. ☐ The failure of the nucleotide and/or amino acid sequence listing to comply with the prescribed requirements prevents a meaningful search from being carried out:
   - ☐ it does not comply with the prescribed standard
   - ☐ it is not in the prescribed machine readable form
4. Further comments:
   see next sheet

Name and mailing address of the International Searching Authority	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL-2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Peter Hedman/EK Telephone no. 08-782 25 00

The indepentent claim 1 suggests a packet switched environment, including a functionality of a presence server. This claim is unclear since it fails to define the matter for which protection is sought (See PCT Art. 6).

Claims 2-36 all depend on claim 1. It is unclear in what context the matter specified in these claims is supposed to be implemented. For this reason, also these claims are unclear.

Consequently, claim 1-36 fail to comply with the prescribed requirements to such an extent that no meaningful search can be carried out (See Art. 17 (2)(a)(ii)).